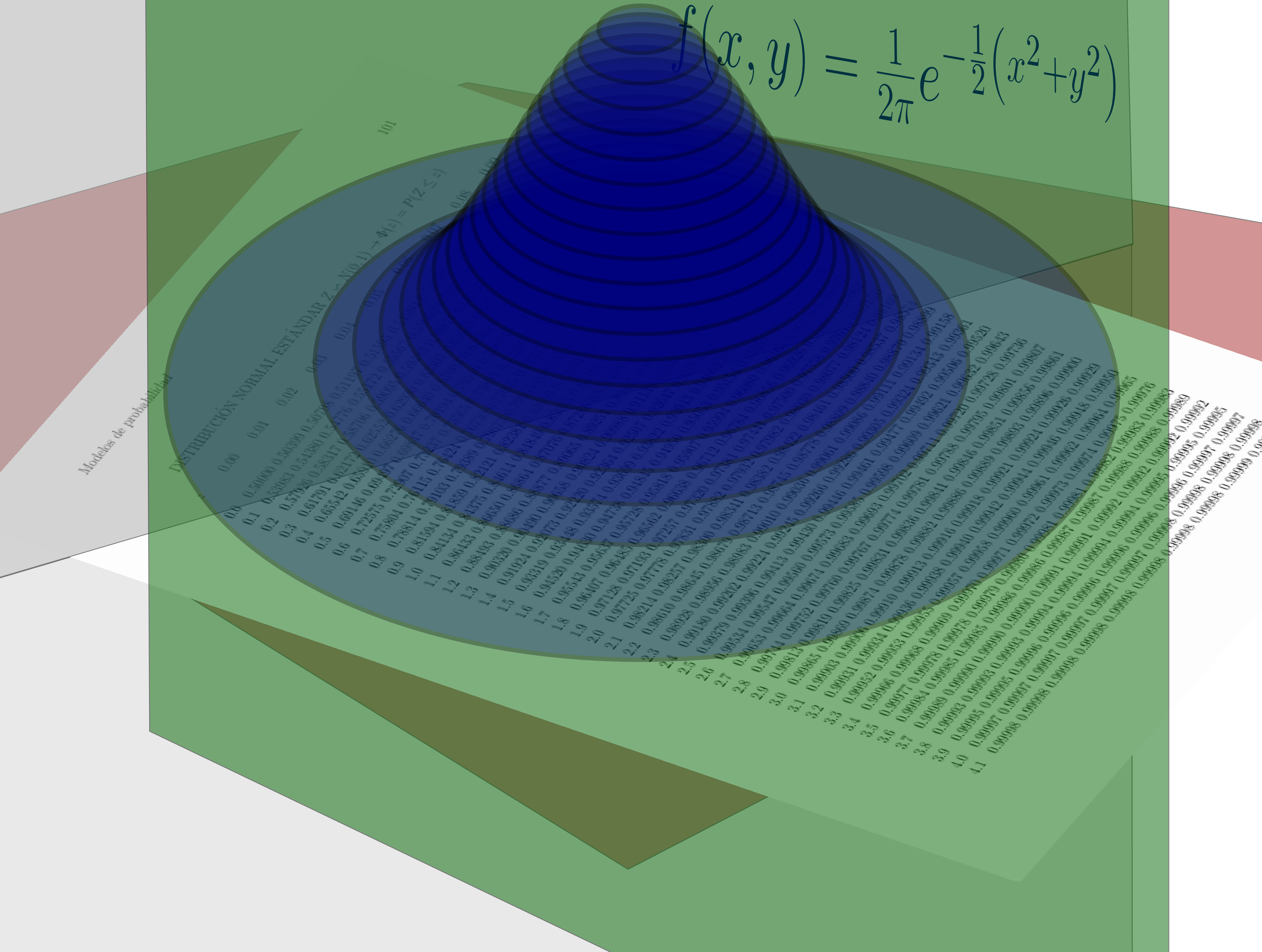
Ejercicios de Probabilidad

Jesús Juan Ruiz



Universidad Politécnica de Madrid

Septiembre, 2018

Ejercicios de Probabilidad

Jesús Juan Ruiz

Universidad Politécnica de Madrid

Septiembre 2018

Índice general

Capítulo 1

Probabilidad elemental

1. Sean A y B dos sucesos no disjuntos cualesquiera de un experimento aleatorio. Indicar cual de las siguientes probabilidades es mayor $P_1 = P(A \cap B \mid A)$ ó $P_2 = P(A \cap B \mid A \cup B)$. Justificar la respuesta.

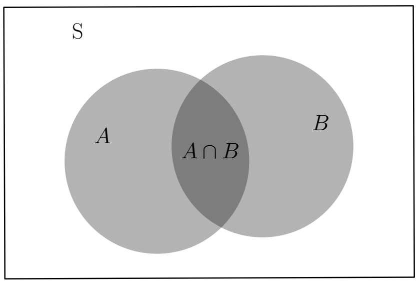


Figura 1.1: Intersección de dos sucesos

Solución:

Mira la figura 1.1 : P_1 es el cociente entre el área de $A \cap B$ y el área de A. P_2 es el cociente entre el área de $A \cap B$ y el área de $A \cup B$. Como el área de $A \cup B$ es mayor que el área de A, es evidente que $P_1 \geq P_2$. Vamos a razonarlo de otra forma: como $(A \cap B) \cap A = A \cap B$, se tiene que

$$P_1 = P(A \cap B \mid A) = \frac{P(A \cap B)}{P(A)} \Longrightarrow P(A \cap B) = P_1 P(A)$$

y como $(A \cap B) \cap (A \cup B) = A \cap B$, se tiene que

$$P_2 = P(A \cap B \mid A \cup B) = \frac{P(A \cap B)}{P(A \cup B)} \Longrightarrow P(A \cap B) = P_2 P(A \cup B)$$

por tanto

$$P_1 P(A) = P_2 P(A \cup B) \Longrightarrow \frac{P_1}{P_2} = \frac{P(A \cup B)}{P(A)} \geq 1$$

pues $P(A \cup B) \geq P(A)$ y por tanto $P_1 \geq P_2$.

• • •

2. Dado dos sucesos A y B, tal que $P(A) > 0$ y $P(B) > 0$ indicar, justificando la respuesta, si son ciertas o no las afirmaciones siguientes:

   a) Si A y B son sucesos mutuamente excluyentes entonces no pueden ser independientes.

   b) Si A y B son independientes entonces no pueden ser mutuamente excluyentes.

   Solución:

   a) Si los sucesos A y B son excluyentes entonces $A \cap B = \emptyset$ y entonces $P(A \cap B) = 0$. Pero nos dicen que $P(A) > 0$ y $P(B) > 0$, entonces $P(A)P(B) > 0$, por tanto si son excluyentes no pueden ser independientes.

   b) Si los sucesos A y B son independientes, entonces $P(A \cap B) = P(A)P(B)$. Si dos sucesos cumplen que $P(A) > 0$ y $P(B) > 0$ entonces $P(A \cap B) > 0$. Pero entonces no pueden ser excluyentes ($P(A \cap B) = 0$). Por tanto si dos sucesos son independientes, no pueden ser excluyentes.

• • •

3. Demostrar que si dos sucesos A y B son independientes, también los son A y $\overline{B}$, y $\overline{A}$ y $\overline{B}$, donde $\overline{A}$ y $\overline{B}$ representan los sucesos complementarios de A y B respectivamente.

   Solución:

   Primero vamos a demostrar que si A y B son sucesos independientes, entonces A y $\overline{B}$ también lo son. El suceso A es la unión de $A \cap B$ y $A \cap \overline{B}$, además los sucesos $A \cap B$ y $A \cap \overline{B}$ son disjuntos. Por tanto:

$$A = (A \cap B) \cup (A \cap \overline{B})$$

$$P(A) = P(A \cap B) + P(A \cap \overline{B})$$

   Como A y B son independientes, $P(A \cap B) = P(A)P(B)$, por tanto

$$P(A) = P(A)P(B) + P(A \cap \overline{B})$$

   Despejando se tiene

$$P(A \cap \overline{B}) = P(A) - P(A)P(B) = P(A)(1 - P(B)) = P(A)P(\overline{B}).$$

   Por tanto A y $\overline{B}$ son independientes.

Si A y $\overline{B}$ son dos sucesos independientes, aplicando la misma propiedad se sabe que $\overline{A}$ y $\overline{B}$ también lo son. (Llama $C = \overline{B}$, si A y C son independientes, según acabamos de demostrar entonces $\overline{A}$ y C son independientes).

Es decir, enlazando la propiedad dos veces se tiene que si A y B son independientes lo son $\overline{A}$ y $\overline{B}$.

Para los que no estén convencidos, vamos a demostrar de otra forma que si A y B son independientes, entonces $\overline{A}$ y $\overline{B}$, también son independientes. Por una de las leyes de Morgan se tiene que $\overline{A \cup B} = \overline{A} \cap \overline{B}$, por tanto

$$\begin{aligned}
P(\overline{A} \cap \overline{B}) &= P(\overline{A \cup B}) \\
&= 1 - P(A \cup B) \\
&= 1 - (P(A) + P(B) - P(A \cap B)) \\
&= 1 - (P(A) + P(B) - P(A)P(B)) \\
&= (1 - P(A))(1 - P(B)) \\
&= P(\overline{A})P(\overline{B}).
\end{aligned}$$

• • •

4. Sean dos sucesos equiprobables e independientes tal que la probabilidad de su unión es 0.75. Obtener la probabilidad de cada uno de ellos.

   Solución:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0{,}75$$

   Por ser independientes, $P(A \cap B) = P(A)P(B)$. Además son equiprobables, llamando $p = P(A) = P(B)$ y sustituyendo en la ecuación anterior:

$$p + p - p^2 = 0{,}75 \Rightarrow p^2 - 2p + \frac{3}{4} = 0$$

   La ecuación de segundo grado que tiene como soluciones

$$p_1 = \frac{2 + \sqrt{4-3}}{2} = 1{,}5$$

   y

$$p_2 = \frac{2 - \sqrt{4-3}}{2} = 0{,}5$$

   La solución p_1 no es válida, por tanto la única solución es $p_2 = 0{,}5$.

• • •

5. Se ha realizado una encuesta entre los estudiantes de grado del MIT (Massachusetts Institute of Technology) para conocer sus preferencias tecnológicas. El 35 % de los entrevistados tienen un iPhone y un IPad, el 80 % tienen al menos uno de estos dispositivos y el 60 % no tiene iPad. Se elige un estudiante al azar, calcula la probabilidad de que:

   a) Disponga de iPhone y no de iPad.
   b) Tenga un iPad pero no un Iphone.
   c) Tenga únicamente uno de los dos dispositivos.
   d) No Disponga de ninguno de los dos dispositivos.

   Solución: Sea el suceso A " el estudiante tiene un iPhone" y el B, "el estudiante tiene un iPad". Nos dan como información: $P(A \cap B) = 0{,}35$, $P(A \cup B) = 0{,}80$ y $P(\overline{B}) = 0{,}60$. Vamos a contestar a las preguntas:

   a) Se quiere calcular $P(A \cap \overline{B})$, teniendo en cuenta que $A = (A \cap B) \cup (A \cap \overline{B})$, y que $(A \cap B)$ y $(A \cap \overline{B})$ son disjuntos,

   $$P(A) = P(A \cap B) + P(A \cap \overline{B}) \Rightarrow P(A \cap \overline{B}) = P(A) - P(A \cap B) \qquad (1.1)$$

   Por otra parte se tiene que

   $$P(A \cup B) = P(A) + P(B) - P(A \cap B) \Rightarrow P(A) = P(A \cup B) - P(B) + P(A \cap B)$$

   Sustituyendo $P(A)$ en (1.1) se tiene que

   $$P(A \cap \overline{B}) = P(A \cup B) - P(B) = 0{,}80 - 0{,}40 = 0{,}40$$

   b) Nos piden $P(\overline{A} \cap B)$, que utilizando $P(B) = P(A \cap B) + P(\overline{A} \cap B)$, se tiene

   $$P(\overline{A} \cap B) = P(B) - P(A \cap B) = 0{,}40 - 0{,}35 = 0{,}05$$

   c) Que sólo tenga uno de los dispositivos es

   $$P(A \cap \overline{B}) + P(\overline{A} \cap B) = 0{,}40 + 0{,}05 = 0{,}45$$

   d) Dados dos sucesos A y B se tienen cuatro opciones que son los cuatro sucesos disjuntos siguientes: $A \cap B$, $A \cap \overline{B}$, $\overline{A} \cap B$ y $\overline{A} \cap \overline{B}$. Llamando a, b, c y d a las probabilidades de los cuatro sucesos, se tiene que $a + b + c + d = 1$. En nuestro caso $a + b + c = 0{,}8$, por tanto $d = 0{,}20$, es decir $P(\overline{A} \cap \overline{B}) = 0{,}2$. En la siguiente tabla se tiene un esquema de los distintos sucesos del ejercicio. Para completar la tabla es suficiente con conocer o poder calcular tres de los cuatro valores a, b, c y d.

	iPAD	No-iPad	
iPhone	a	b	$a+b$
No-iPhone	c	d	$c+d$
	$a+c$	$b+d$	1

$\Rightarrow$

	iPAD	No-iPad	
iPhone	0,35	0,40	0,75
No-iPhone	0,05	0,20	0,25
	0,40	0,60	1

• • •

6. Una encuesta realizada entre inmigrantes proporciona la siguiente información: el 80 % de los jóvenes entre 18 y 25 años no tiene trabajo, el 75 % de los jóvenes en esa edad no están matriculados en ningún centro educativo (no estudian). Si se toma un joven al azar de la muestra,¿cuál es la probabilidad de que ni estudie ni trabaje? (Otra forma de hacer la misma pregunta, ¿cuál es la proporción de jóvenes de la encuesta que ni estudian ni trabajan?

   Solución: Si tomamos un individuo al azar de la muestra la probabilidad de que no tenga trabajo es 0.80 y la probabilidad de que no estudie 0.75. Llamando A al suceso "no trabajanz B al suceso "no estudian", nos piden obtener $P(A \cap B)$. Con la información que se tiene no se puede calcular exáctamente, pero se puede acotar. Teniendo en cuenta que $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ y sabiendo que tiene $P(A \cup B) \leq 1$, tenemos que

   $$0{,}80 + 0{,}75 - P(A \cap B) \leq 1 \qquad \Rightarrow P(A \cap B) \geq 0{,}55$$

   Por otro lado, la intersección no puede ser mayor que cualquiera de los dos sucesos, por tanto $P(A \cap B) \leq P(B) = 0{,}75$. Es decir, se puede afirmar que la proporción de personas en la muestra que ni estudian ni trabajan cumple $0{,}55 \leq P(A \cap B) \leq 0{,}75$

• • •

7. Demostrar que para cualquier par de sucesos, A_1, A_2 se cumple que

   $$P(A_1 \cap A_2) \geq P(A_1) + P(A_2) - 1,$$

   demostrar que la extensión a n sucesos es $P(A_1 \cap A_2 \cap \cdots \cap A_n) \geq \sum_{i=1}^{n} P(A_i) - (n-1)$.

   Solución:

   La probabilidad de la unión de dos sucesos cumple que

   $$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) \leq 1$$

   despejando $P(A_1 \cap A_2)$, se obtiene

   $$P(A_1 \cap A_2) \geq P(A_1) + P(A_2) - 1.$$

   Que es la primera pregunta del ejercicio.

   Para demostrar la segunda, aplicaremos indución. Se cumple para $n = 2$ como hemos demostrado. Suponemos que se cumple para $n - 1$, es decir que

   $$P(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) \geq \sum_{i=1}^{n-1} P(A_i) - (n - 2). \tag{1.2}$$

Llamamos $B_{n-1} = A_1 \cap A_2 \cap \cdots \cap A_{n-1}$, aplicamos la propiedad para $n = 2$, con B_{n-1} y A_n,

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(B_{n-1} \cap A_n) \geq P(B_{n-1}) + P(A_n) - 1$$

Sustituyendo el resultado de $P(B_{n-1})$ de la ecuación 1.2 se tiene que

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) \geq \sum_{i=1}^{n} P(A_i) - (n-1).$$

Veamos un ejemplo: si el porcentaje de hombres rubios en un país es del 90 % y el porcentaje de hombres con ojos azules es 85 %, según lo demostrado, en el país hay al menos 75 % de hombres rubios y con ojos azules $(0{,}90 + 0{,}85 - 1 = 0{,}75)$. Si además nos dicen que hay un 95 % con carnet de conducir, pues podremos afirmar que al menos el 70 % son hombres rubios con ojos azules y carnet de conducir $(0{,}90 + 0{,}85 + 0{,}95 - 2 = 0{,}70)$.

• • •

8. La probabilidad de que un componente de una máquina se averíe antes de 100 horas es 0.01. La máquina tiene 50 componentes; calcular la probabilidad de avería de la máquina antes de 100 horas en los casos siguientes:

    a) La máquina se avería cuando lo hace uno o más componentes.

    b) La máquina se avería cuando fallan dos o más componentes.

    c) La máquina sólo se avería cuando lo hacen todos los componentes.

    Solución:

    Vamos a llama F_i al suceso "*el componente i funcionaz* $\overline{F}_i$ al contrario. Para poder resolverlo, supondremos que el fallo de los componentes son sucesos independientes. La probabilidad de que un componente falle es $p = 0{,}01$.

    - Para que funcione la máquina, tienen que funcionar todos los componentes

    $$P(F_1 \cap F_2 \cap \cdots \cap F_{50}) = P(F_1)P(F_2)\cdots P(F_{50}) = (1-p)^{50} = 0{,}99^{50} = 0{,}605$$

    - Para que funcione la máquina tiene que ocurrir que: funcionen todos, o que funcionen todos menos el primero, o que funcionen todos menos el segundo, ..., que funcionen todos menos el quincuagésimo. La probabilidad de que funcionen todos menos el primero es $p(1-p)^{49}$, la probabilidad de que funcionen todos menos el segundo es la misma $p(1-p)^{49}$ y as hasta el último caso, que funcionen todos menos el quincuagésimo que tiene probabilidad $p(1-p)^{49}$. Ver el esquema en la figura 1.2. La probabilidad de que la áquina funciones es

    $$(1-p)^{50} + 50p(1-p)^{49} = 0{,}9106$$

$$\begin{array}{ll}
\text{Todos Funcionan} & F_1 \cap F_2 \cap F_3 \cap \cdots \cap F_{50} \to (1-p)^{50} \\
\left.\begin{array}{ll}
\text{Falla el 1, los demás funcionan} & \overline{F}_1 \cap F_2 \cap F_3 \cap \cdots \cap F_{50} \to p(1-p)^{49} \\
\text{Falla el 2, los demás funcionan} & F_1 \cap \overline{F}_2 \cap F_3 \cap \cdots \cap F_{50} \to p(1-p)^{49} \\
 & \ldots \\
\text{Falla el 50, los demás funcionan} & F_1 \cap F_2 \cap F_3 \cap \cdots \cap \overline{F}_{50} \to p(1-p)^{49}
\end{array}\right) & 50p(1-p)^{49}
\end{array}$$

Figura 1.2: La máquina funciona si solo falla un componente

- Para que falle la máquina, tienen que fallar todos los componentes

$$P(\overline{F}_1 \cap \overline{F}_2 \cap \cdots \cap \overline{F}_{50}) = P(\overline{F}_1)P(\overline{F}_2)\cdots P(\overline{F}_{50}) = p^{50} = 1{,}0 \times 10^{-100}$$

• • •

9. Tres fichas están marcadas con las letras A, B, C, respectivamente, y una cuarta con las tres ABC. Se meten en una urna y se toma una de ellas al azar. Se pregunta si los tres sucesos consistentes en la presencia de la letra A, la letra B o la C sobre la ficha son o no independientes.

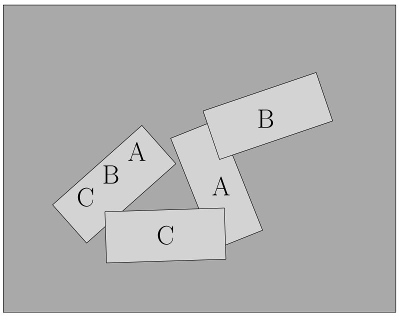


Figura 1.3: Cuatro fichas con letras A, B, C y ABC

Solución: Sea X_A, X_B y X_C el suceso la tarjeta extraída tiene la letra A, B, o C. De manera que $X_A = \{A, ABC\}$, $X_B = \{B, ABC\}$, y $X_C = \{C, ABC\}$. Está claro que $P(X_A) = P(X_B) = P(X_C) = 2/4$. Como $X_A \cap X_B = \{ABC\}$, se tiene que $P(X_A \cap X_B) = \dfrac{1}{4}$, y observamos que se cumple que X_A y X_B son sucesos independientes, pues

$$P(X_A \cap X_B) = P(X_A)P(X_B) = \frac{1}{4}.$$

Con el mismo razonamiento se prueba que X_A y X_C son dos sucesos independientes, y tambien X_B y X_C son independientes.

Sin embargo, los tres, no son independientes pues $X_A \cap X_B \cap X_C = \{ABC\}$, y $P(X_A \cap X_B \cap X_C) = \frac{1}{4}$, y sin embargo

$$P(X_A)P(X_B)P(X_C) = (\frac{1}{2})^3 \neq P(X_A \cap X_B \cap X_C) = \frac{1}{4}.$$

• • •

10. En un campeonato de tenis usted tiene la opción de escoger la secuencia de partidos $A - B - A$ o la $B - A - B$, donde A y B indican sus oponentes. Para clasificarse debe usted ganar dos partidos consecutivos. El jugador A es mejor que el B. ¿ Qué secuencia será preferida?

    Solución: Sea S_1 la secuencia $A-B-A$ y S_2 la secuencia $B-A-B$. La probabilidad de ganar a A se denota como P_A y P_B la probabilidad de ganar a B. Nos dicen que $P_A < P_B$. Denominamos G_A y G_B los sucesos ganar a A y a B, y $\overline{G}_A$, $\overline{G}_B$ sus complementarios. La probabilidad de clasificarse para cada secuencia es:

$$\begin{aligned}
P(\text{Clasificarse}|S_1) &= P(G_A \cap G_B \cap G_A) + P(G_A \cap G_B \cap \overline{G}_A) + P(\overline{G}_A \cap G_B \cap G_A) \\
&= P_A^2 P_B + 2P_A P_B(1 - P_B) \\
&= P_A P_B(2 - P_A)
\end{aligned}$$

$$\begin{aligned}
P(\text{Clasificarse}|S_2) &= P(G_B \cap G_A \cap G_B) + P(G_B \cap G_A \cap \overline{G}_B) + P(\overline{G}_B \cap G_A \cap G_B) \\
&= P_A P_B^2 + 2P_A P_B(1 - P_B) \\
&= P_A P_B(2 - P_B)
\end{aligned}$$

    Como $(2 - P_A) > (2 - P_B)$, es mejor la secuencia $A - B - A$, aunque A sea mejor que B. Se ha supuesto independencia.

• • •

11. Un jurado de tres miembros que decide por mayoría tiene dos miembros que deciden independientemente el veredicto correcto con probabilidad p y el tercero lanza una moneda. Si un juez tiene probabilidad p de acertar, ¿cuál de los dos sistemas tiene mayor probabilidad de acertar?

Solución: Llamamos M_i al suceso "el miembro i del jurado acierta" siendo $i = A, B$, o C. Si deciden por mayoría, el jurado acierta si aciertan dos o más de sus miembros.

$$P(M_A \cap M_B \cap M_C) = p \times p \times \frac{1}{2} \qquad \text{aciertan los 3}$$
$$P(M_A \cap M_B \cap \overline{M}_C) = p \times p \times \frac{1}{2} \qquad \text{aciertan A y B; C se equivoca}$$
$$P(M_A \cap \overline{M}_B \cap M_C) = p \times (1-p) \times \frac{1}{2} \qquad \text{aciertan A y C; B se equivoca}$$
$$P(\overline{M}_A \cap M_B \cap M_C) = (1-p) \times p \times \frac{1}{2} \qquad \text{aciertan B y C; A se equivoca}$$

Por tanto, la suma es la probabilidad de que el jurado acierte

$$\frac{1}{2}p^2 + \frac{1}{2}p^2 + \frac{1}{2}p(1-p) + \frac{1}{2}p(1-p) = p^2 + p(1-p) = p$$

Este jurado tan *particular* tiene la misma probabilidad de acertar que el juez.

• • •

12. Una caja contiene r bolas rojas y b bolas blancas. Se elige una bola al azar de la caja (es decir todas las bolas tienen las misma probabilidad de ser elegidas), y después se extrae otra al azar entre las restantes de la caja. Calcular la probabilidad de

    a) Las dos bolas sean rojas

    b) Las dos bolas sean blancas

    c) La primera bola sea roja y la segunda blanca

    d) Las dos bolas sean de distinto color

    Solución: Voy a resolver el problema por combinatoria

    a) En la caja hay $r+b$ bolas, cada bola tiene la misma probabilidad de ser elegida en la primera extracción. La primera extración da lugar a $r+b$ posibilidaes (equiprobables). Piensa que las bolas están numeradas desde 1 hasta $r+b$ y olvídate del color de la bola (ver figura 1.4). Una vez elegida la primera bola, como es sin reposición, para la segunda teneremos $r+b-1$ opciones, por tanto en total existen $(r+b)(r+b-1)$ formas distintas de extraer dos bolas sin reposición de la caja. Todas equiprobables. (Ver figura 1.4 a).

    Ahora tenemos que contar cuantas de las anteriores configuraciones están fromadas por bolas rojas. O lo que es lo mismo, de cuantas formas se pueden extraer dos bolas rojas de la caja. Aplicamos el mismo razonamiento que antes: para la primera elección se tienen r posibilidades y para la segunda $r-1$, por consiguiente hay $r(r-1)$ resultados con dos bolas rojas.

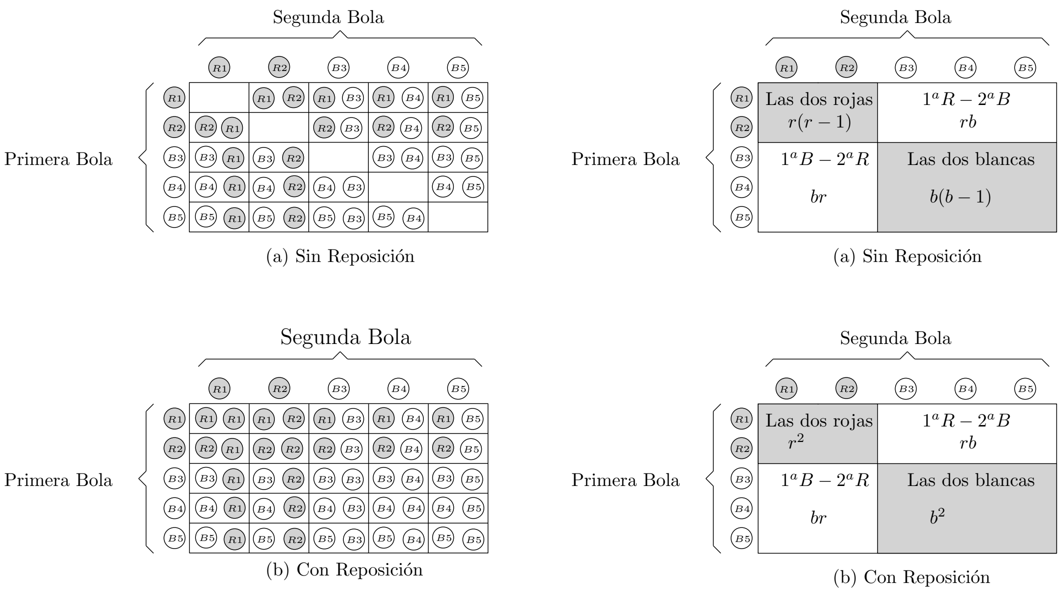


Figura 1.4: Caja con r bolas rojas y b blancas

Como todos los sucesos son equiprobables, la probabilidad pedida es el cociente entre el número de casos favorables y casos posibles:

$$P(\text{“Dos bolas rojas”}) = \frac{r(r-1)}{(r+b)(r+b-1)}$$

b) El número de posibles resultados equiprobables son los mismos, $(r+b)(r+b-1)$. Los resultados que tienen dos bolas blancas son $b(b-1)$, por tanto,

$$P(\text{“Dos bolas blancas”}) = \frac{b(b-1)}{(r+b)(r+b-1)}$$

c) Las configuraciones con la primera bola roja y la segunda blanca son rb, por tanto

$$P(\text{“La primera roja y la segunda blanca”}) = \frac{rb}{(r+b)(r+b-1)}$$

d) Las configuraciones con dos bolas de color diferente son las calculadas en el apartado anterior rb, más las que cumplen que la primera es blanca y la segunda roja, que son br, por tanto el número de configuraciones con dos bolas de colores distintos $2rb$, y

$$P(\text{“Las dos bolas diferentes”}) = \frac{2rb}{(r+b)(r+b-1)}$$

Es muy fcil comprobar (ver figura 1.4 a) que el número total de configuraciones cumple $(r+b)(r+b-1) = r(r-1) + b(b-1) + 2rb$.

• • •

13. Una caja contiene r bolas rojas y b bolas blancas. Se elige una bola al azar de la caja (es decir todas las bolas tienen las misma probabilidad de ser elegidas), después se devuelve a la caja, se mezcla y se extrae otra al azar. Calcular la probabilidad de

    a) Las dos bolas sean rojas
    b) Las dos bolas sean blancas
    c) La primera bola sea roja y la segunda blanca
    d) Las dos bolas sean de distinto color

    Solución: Razonando de forma similar al ejercicio anterior (Ver figura 1.4 b).

    a) En la primera extracción se tienen $r+b$ posibles resultados, una vez seleccionada la primera bola, para la segunda se tienen otra vez $r+b$ opciones, por tanto existen $(r+b)^2$ posibles resultados. Todos equiprobables. De esas configuraciones hay r^2 con las dos bolas rojas, la primera elección se tienen r posibilidades y para la segunda otras r,

    $$P(\text{“Dos bolas rojas”}) = \frac{r^2}{(r+b)^2}$$

    b) De igual forma

    $$P(\text{“Dos bolas blancas”}) = \frac{b^2}{(r+b)^2}$$

    c) Las configuraciones con la primera bola roja y la segunda blanca son rb, por tanto

    $$P(\text{“La primera roja y la segunda blanca”}) = \frac{rb}{(r+b)^2}$$

    d) El número de configuraciones con dos bolas de colores distintos $2rb$, y

    $$P(\text{“Las dos bolas diferentes”}) = \frac{2rb}{(r+b)^2}$$

    Ahora se tiene que el número total de configuraciones cumple $(r+b)^2 = r^2 + b^2 + 2rb$.

• • •

14. Una urna contiene 3 bolas azules, 2 blancas y 5 rojas. Si se extraen 6 bolas con reposición, calcular la probabilidad de obtener 2 bolas de cada color.

    Solución:

    La probabilidad de obtener (exáctamente y es este orden) la secuencia $R_1 \cap R_2 \cap A_3 \cap A_4 \cap B_5 \cap B_6$ es:

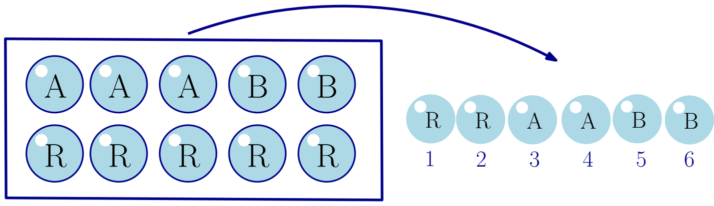


Figura 1.5: Urna con 10 bolas

$$
\begin{aligned}
P(R_1 \cap R_2 \cap A_3 \cap A_4 \cap B_5 \cap B_6) &= P(R_1)P(R_2)P(A_3)P(A_4)P(B_5)P(B_6) \\
&= \left(\frac{5}{10}\right)^2 \left(\frac{3}{10}\right)^2 \left(\frac{2}{10}\right)^2,
\end{aligned}
$$

pues la extración es con reposición y todos los sucesos son independientes. Hay muchas configuraciones (ordenaciones) distintas de 6 bolas con dos bolas de cada color. Por ejemplo $R_1 \cap A_2 \cap B_3 \cap R_4 \cap A_5 \cap B_6$, la probabilidad de exáctamente este resultado es la misma $\left(\frac{5}{10}\right)^2 \left(\frac{3}{10}\right)^2 \left(\frac{2}{10}\right)^2$. Para completar el ejercicio es preciso "contar.[el] número de ordenaciones distintas de las letras $RRAABB$. Si se dispone de n letras distintas, el número de reordenaciones de esas n letras es $n!$. Si de las n letras hay r distintas, n_1 del tipo 1, n_2 del tipo 2, y ass sucesivamente hasta n_r del tipo r, donde $n_1 + n_2 + \cdots + n_r = n$, el número de configuraciones distintas es

$$\frac{n!}{n_1!n_2!\cdots n_r!}.$$

En combinatoria este caso se denomina permutaciones con repetición. En nuestro ejemplo tenemos 6 letras, R se repite 2 veces, A se repite 2 veces y B se repite 2 veces, por tanto el número de reordenaciones distintas es

$$\frac{6!}{2!\,2!\,2!} = 90.$$

La probabilidad de cada una es la misma $\left(\frac{5}{10}\right)^2 \left(\frac{3}{10}\right)^2 \left(\frac{2}{10}\right)^2$, por tanto la probabilidad pedida es

$$\frac{6!}{2!\,2!\,2!} \left(\frac{5}{10}\right)^2 \left(\frac{3}{10}\right)^2 \left(\frac{2}{10}\right)^2 = 0{,}081$$

Cuando se disponen de n elementos y solo hay de dos tipos, en nuestro caso dos colores, por ejemplo blancas y negras, se habla de un problema binomial. Cuando hay tres tipos diferentes se habla de una distribución trinomial y en general se denomina problema multinomial.

Otra extensión al problema es el caso en el que la extracción se hace sin reposición. La probabilidad en este caso es

$$\frac{\binom{5}{2}\binom{3}{2}\binom{2}{2}}{\binom{10}{6}} = 0{,}14286$$

La generalización a otras situaciones es directa.

• • •

15. El sorteo de la lotería primitiva consiste en acertar los seis números extraídos al azar y sin reposicin de una bombo que contiene 49 bolas numeradas del 1 al 49. Un participante ha realizado una apuesta múltiple y ha tachado 8 números. Calcular la probabilidad de acertar 6, 5, 4, 3, 2, 1 o ninguno de los números premiados. Si juega todos los días del año (365 sorteos) la misma apuesta ¿cuál es la probabilidad de que no acierte más de cuatro en ningún sorteo del año?

    Solución:

    Supongamos que hemos elegido (tachado) los números $B = \{1, 3, 5, 7, 11, 13, 17, 23\}$. Denominamos N el conjunto complementario, formado por los 41 números no elegidos $N = \{2, 4, 6, 8, 9, ..., 49\}$. El bombo contiene dos tipos de bolas, las 8 que hemos elegido y las 41 que no hemos elegido. Podemos visualizar el problema imaginando que los 8 números elegidos son bolas de color blanco y las restantes 41 negras, correspondientes al conjunto N (ver figura 1.6).

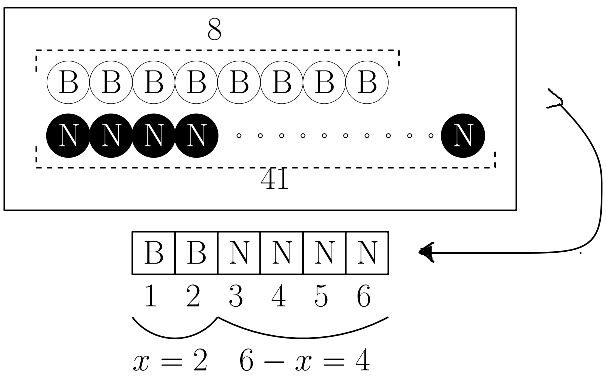


Figura 1.6: Lotería Primitiva: Hemos elegido 8 números, ¿ cuál es la probabilidad de acertar 2 al sacar 6?

En el sorteo se extraen 6 bolas (sin reposición). Cualquier combinación de 6 números tiene la misma probabilidad de ser extraída. El número de combinaciones posibles son $\binom{49}{6} = 13\,983\,816$. El número de combinaciones que tienen x bolas blancas y $6 - x$ bolas negras es

$$\binom{8}{x}\binom{41}{6-x}$$

y por tanto la probabilidad de acertar x es

$$P(\text{ Acertar } x) = \frac{\binom{8}{x}\binom{41}{6-x}}{\binom{49}{6}}, \text{ con } x = 0, 1, 2, 3, 4, 5, 6$$

Tabla 1.1: Lotería Primitiva

Aciertos	Probabilidad
0	0.321542
1	0.428723
2	0.202774
3	0.042689
4	0.004105
5	0.0001641898
6	0.0000020023

Tabla 1.2: Probabilidades de acertar $x = 0, 1, ..., 6$ con una apuesta de 8 números

En la tabla 1.2 se proporcionan los valores. Se observa que la probabilidad de acertar 5 o 6 es muy baja

$$P(\text{Acertar 5 o 6}) = 0{,}0001641898 + 0{,}0000020023 = 0{,}000166192$$

y la probabilidad de acertar menos de 5 es muy alta: $1 - 0{,}000166192 = 0{,}9998338$. La probabilidad de acertar menos de 5 en los 365 sorteos del año es $0{,}9998338^{365} = 0{,}9411$. Es decir, la probabilidad de tener un buen premio (acertar 5 o 6) al menos una vez durante el año es $1 - 0{,}9411 = 0{,}0588$.

• • •

16. Si la probabilidad de que un interruptor cualquiera de la figura 1.7 se encuentre cerrado (es decir, pase corriente) es p, calcular la probabilidad de que pase corriente de A a B para las tres configuraciones de la figura (1.7).

Solución:

Antes analizar los circuitos del ejercicio vamos a resolver primero dos configuraciones muy simples(1.8).

a) La primera tiene dos interruptores en serie (ver Figura 1.8a). Los dos interruptores tienen la misma probabilidad p de estar cerrados y actúan independientemente. Para que pase corriente de A a B tienen que estar cerrado los dos, por tanto la probabilidad de que circule corriente por el circuito es p^2.

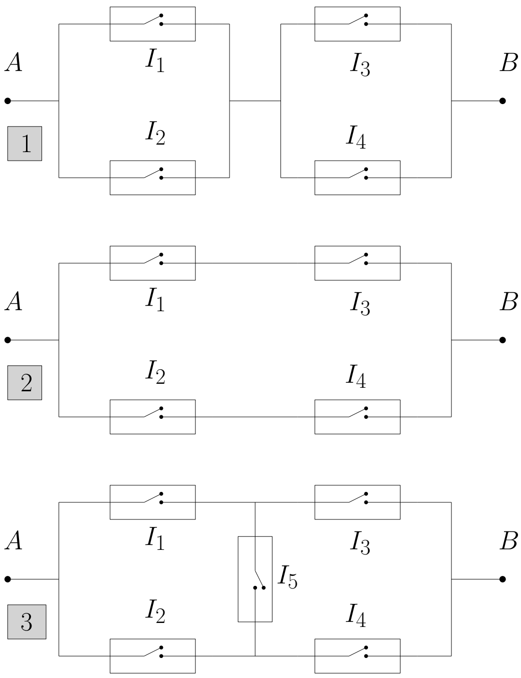
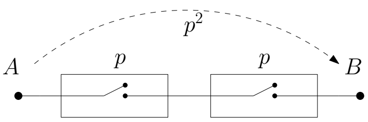
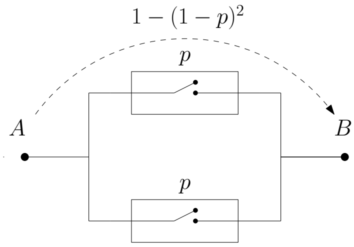


Figura 1.7: Tres circuitos



(a) Dos interruptores en serie



(b) Dos interruptores en paralelo

Figura 1.8: Configuraciones básicas

b) La segunda tiene los dos interruptores en paralelo (ver Figura 1.8b). Para que no pase corriente de A a B, los dos tienen que estar abiertos. La probabilidad de que los dos estén abiertos es $(1-p)^2$, el suceso contrario es que al menos uno esté cerrado y **pase corriente**, $1-(1-p)^2$.

Con los dos resultados anteriores podemos resolver cualquier tipo de circuito. Vamos a aplicarlo a las tres configuraciones del ejercicio.

Configuración 1 Podemos dividir el circuito en dos bloques en serie G_1 y G_2. Para que pase corriente de A a B tienen que pasar corriente por G_1 y por G_2. Denominamos A_1 al suceso "pasa corriente por el grupo G_1z A_2 al suceso "pasa corriente por el grupo G_2". Como hemos visto $P(A_1) = P(A_2) = 1-(1-p)^2$. Llamando P_1 a la probabilidad de que pase corriente de A a B se tiene que

$$P_1 = P(A_1 \cap A_2) = P(A_1)P(A_2) = \left[1-(1-p)^2\right]^2$$

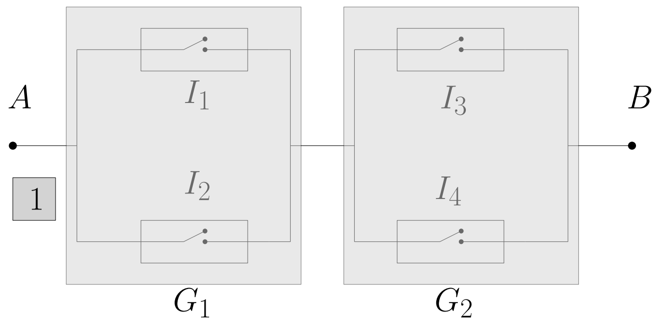


Figura 1.9: Grupos en serie

Configuración 2 En la configuración 2 los grupos ahora están en paralelo. Para que no pase corriente de A a B, tienen que estar abierto los dos grupos. Denominamos B_1 al suceso "pasa corriente por el grupo G_1 " y B_2 al suceso "pasa corriente por el grupo G_2 ". Como hemos visto $P(\overline{B}_1) = P(\overline{B}_2) = 1 - p^2$. Si P_2 a la probabilidad de que pase corriente de A a B, $1 - P_2$ es la probabilidad de que no pase corriente

$$1 - P_2 = P(\overline{B_1} \cap \overline{B_2}) = P(\overline{B_1})P(\overline{B_2}) = (1 - p^2)(1 - p^2) = (1 - p^2)^2$$

La probabilidad de que pase corriente es

$$P_2 = 1 - (1 - p^2)^2$$

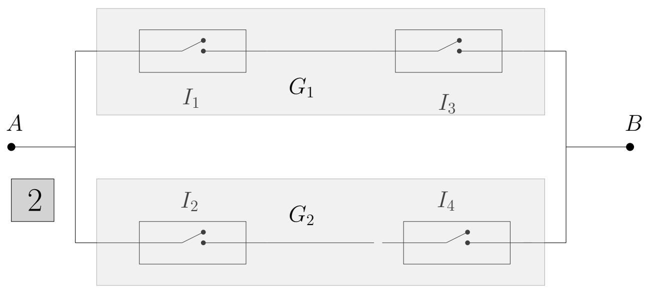


Figura 1.10: Grupos en paralelo

Configuración 3 Es la más complicada, se simplifica si resolvemos el ejercicio considerando dos opciones: La primera opción, que el interruptor I_5 esté cerrado. En este caso, el circuito se corresponde a la configuración 1. La segunda opción es que el interruptor I_5 esté abierto, que es la configuración 2. Llamando C al suceso "pasa corriente de A a B en la configuración 3ž llamando I_5 e $\overline{I_5}$ a los sucesos "pasa corrientež "no pasa corriente"por el interruptor I_5, se puede escribir

$$
\begin{aligned}
P(C) &= P(C \cap I_5) + P(C \cap \overline{I_5}) \\
&= P(C|I_5)P(I_5) + P(C|\overline{I_5})P(\overline{I_5}) \\
&= \left[1 - (1-p)^2\right]^2 \times p + \left[1 - (1-p^2)^2\right] \times (1-p) \\
&= p^2\left(2p^3 - 5p^2 + 2p + 2\right)
\end{aligned}
$$

En la figura (1.11) se muestra las gráficas de la probabilidad de que pase corriente de A a B para las tres configuraciones en función de p. Se aprecia como la probabilidad es siempre mayor para la configuración 1 y menor para la configuración 2, la gráfica de la configuración 3 se encuentra entre las otras dos.

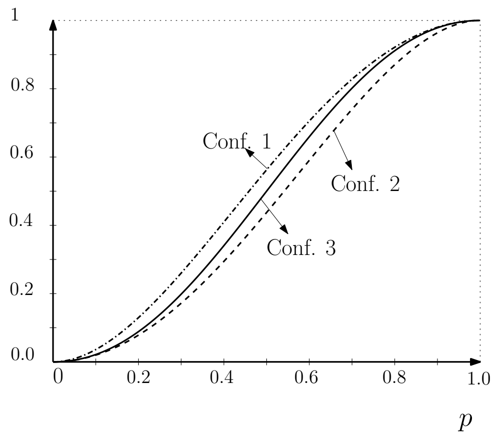


Figura 1.11: Probabilidad de que pase corriente de A a B en las tres configuraciones

• • •

17. Un matrimonio tiene dos hijos y se sabe que uno de ellos es varón. Aceptando que la probabilidad de varón y hembra es la misma. ¿ Cuál es la probabilidad de que los dos sean varones?

    Solución: Es un ejercicio con trampa. Uno puede razonar de este modo: ya que uno de los hijos es varón, solo queda un hijo que mirar; la posibilidad de que ese hijo sea un varón es 50 %; por tanto la probabilidad de que ambos hijos sean varones es del 50 %.

    El razonamiento no es correcto.¿ Por qué?. El enunciado dice que un hijo es varón, pero no dice si es el primero o el segundo y eso cambia las cosas. Partimos del espacio muestral con los cuatro sucesos equiprobables:

$$S = \{V_1 \cap V_2, V_1 \cap M_2, M_1 \cap V_2, M_1 \cap M_2\}$$

El suceso $A =$ "uno es varón", es el subconjuto de S formado por los elementos

$$A = \{V_1 \cap V_2, V_1 \cap M_2, M_1 \cap V_2\}.$$

Los tres sucesos que forman A son equiprobables, por tanto $P(V_1 \cap V_2) = 1/3$.

Otra forma de hacerlo es calcular $P(V_1 \cap V_2|A)$ utilizando el espacio muestral inicial S,

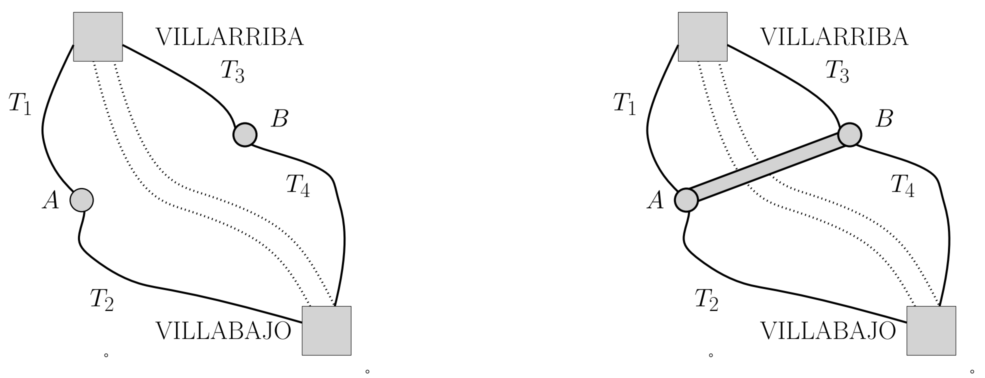
$$P(V_1 \cap V_2|A) = \frac{P(V_1 \cap V_2)}{P(A)} = \frac{1/4}{3/4} = \frac{1}{3}.$$

Si la pregunta hubiera sido *calcular la probabilidad de dos varones sabiendo que el hijo mayor es varón*, entonces $A = \{V_1 \cap V_2, V_1 \cap M_2\}$, y la probabilidad sería del 50 %.

Si ya lo tienes claro resuelve la siguiente variante: Una familia tiene dos hijos y se sabe que uno se llama Juan, ¿Cuál es la probabilidad de que los dos sean varones? La respuesta es 50 %. ¿Por qué?

• • •

18. Dos pueblos de alta montaña están comunicados por cuatro tramos de carretera según se muestra en la Figura 1.12a. Los días de invierno debido a la nieve es frecuente que las carreteras estén cortadas. Cuando nieva, la probabilidad de que cualquier tramo sea transitable es 0.8 con independencia de lo que ocurra con el resto. ¿Cuál es la probabilidad de que un vehículo pueda viajar de Villarriba a Villabajo un día con nieve? Para mejorar las comunicaciones se piensa construir un nuevo tramo de carretera entre los puntos A y B. Cuando nieva la probabilidad de accesibilidad es también 0.8 con independencia del resto. ¿Cuál es la probabilidad de que una persona pueda hacer el viaje anterior con esta nueva configuración 1.12b?



(a) Sin puente (b) Con puente

Figura 1.12: Red de carreteras entre Villarriba y Villabajo

Solución: Este ejercicio es igual al ejercicio de los interruptores (16). Aprovechamos para resolverlo de otra forma. Vamos a denominar Va a Villarriba y Vo a Villabajo.

Sea A el suceso "*se puede viajar de* Va *a* Vo *por* Az B en suceso "*se puede viajar de* Va *a* Vo *por* B". Llamando p a la probabilidad de que se pueda transitar por un tramo ($p = 0{,}8$) se tiene que

$$P(A) = p^2,\ P(\overline{A}) = 1 - p^2$$
$$P(B) = p^2,\ P(\overline{B}) = 1 - p^2$$

Con la información anterior se pueden dar las cuatro opciones que se muestran en la tabla 1.3.

Caso	Suceso	Probabilidad			Pasa
1	$A \cap B$	$P(A \cap B)$	$= P(A)P(B)$	$= p^4$	SI
2	$\overline{A} \cap B$	$P(\overline{A} \cap B)$	$= P(\overline{A})P(B)$	$= (1 - p^2)p^2$	SI
3	$A \cap \overline{B}$	$P(A \cap \overline{B})$	$= P(A)P(\overline{B})$	$= p^2(1 - p^2)$	SI
4	$\overline{A} \cap \overline{B}$	$P(\overline{A} \cap \overline{B})$	$= P(\overline{A})P(\overline{B})$	$= (1 - p^2)^2$	NO

Tabla 1.3: Cuatro posibilidades entre Villarriba y Villabajo

De las cuatro posibilidades, Va y Vo estn comunicados en tres caso y no lo están sólo en el caso $\overline{A} \cap \overline{B}$, por tanto la probabilidad de pasa de Va a Vo es

$$P(\text{Pasar de } Va \text{ a } Vo\ ) = p^4 + 2p^2(1 - p^2) = 1 - (1 - p^2)^2 = 0{,}8704$$

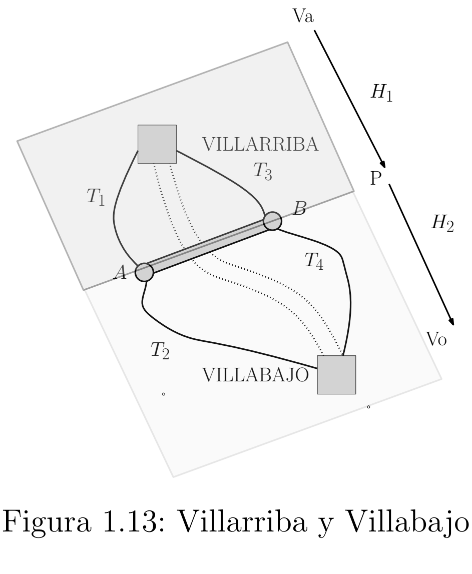


Figura 1.13: Villarriba y Villabajo

Se coloca un puente que va de A a B. Si el puente no es transitable la probabilidad de pasar se Va a Vo es la calculada previamente, $1 - (1 - p^2)^2 = 0{,}8704$.

Analicemos el caso en el que el puente está disponible. Dividimos el trayecto en dos bloques, ver figura 1.13. Vamos a llamar H_1 al suceso se puede pasar de Va a P. Y H_2

al suceso, se puede pasar de P a Vo. Para calculat $P(H_1)$, es más fácil calcular $P(\overline{H}_1)$: para no poder viajar de Va a P, los dos tramos T_1 y T_3 deben estar bloqueados, por tanto $P(\overline{H}_1) = (1-p)^2$, y el suceso contrario tiene probabilidad $P(H_1) = 1-(1-p)^2$. Lo mismo para el bloque que va de P a Vo. Así las probabilidades son:

$$P(\overline{H}_1) = (1-p)^2, \ P(H_1) = 1-(1-p)^2$$
$$P(\overline{H}_2) = (1-p)^2, \ P(H_2) = 1-(1-p)^2$$

Por tanto si el puente está transitable, la probabilidad de pasar de Va a Vo es

$$P(H_1 \cap H_2) = P(H_1)P(H_2) = (1-(1-p)^2)^2 = 0{,}9216$$

.

Si el puente estuviera disponible al 100 %, la probabilidad de comunicación aumenta hasta 0.9216. Pero a veces el puente, también se puede bloquear. Llamando U al suceso "el puente está transitable", $\overline{U}$ "el puente no está transitable", y llamando V al suceso: "Viajar de Va a Vo", se tiene

$$\begin{aligned} P(V) &= P(V \cap U) + P(V \cap \overline{U}) \\ &= P(V|U)P(U) + P(V|\overline{U})P(\overline{U}) \\ &= 0{,}9216 \times 0{,}8 + 0{,}8704 \times 0{,}2 \\ &= 0{,}91008 \end{aligned}$$

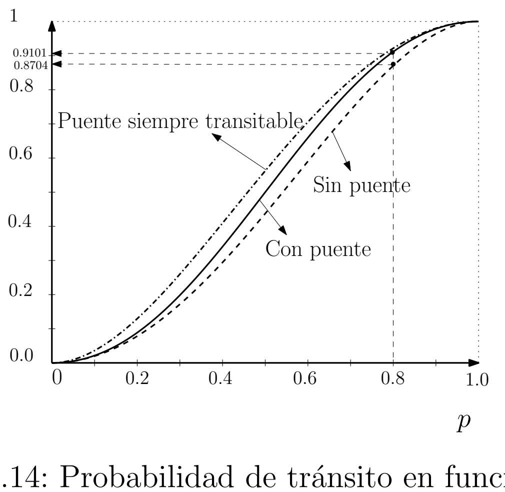


Figura 1.14: Probabilidad de tránsito en función de p

El la figura 1.14 se muestra la probabilidad de poder viajar de Villarriba a Villabajo en función de p. Se muestra las probabilidades para $p = 0{,}8$. Se aprecia que colocar el puente no mejora sustancialmente la comunicación, incluso aunque el puente funcionara el 100 % del tiempo.

• • •

19. Dos estudiantes A y B están matriculados en un curso. A asiste a las clases el 80 % de los días y B el 60 %, siendo las asistencias de ambos independientes. Si exactamente uno de los dos está en clase un día concreto, ¿cuál es la probabilidad de que sea A?

    Solución:

    Llamando A al suceso "A está en clase", B al suceso "B está en clase" y $\overline{A}$ y $\overline{B}$ a sus complementarios, el espacio muestral S se puede descomponer en los siguientes sucesos disjuntos:
    $$A \cap B, \overline{A} \cap B, A \cap \overline{B}, \overline{A} \cap \overline{B}$$

    El suceso "exactamente uno está en clase" está formado por $U = \{\overline{A} \cap B, A \cap \overline{B}\}$. Teniendo en cuenta que $A \cap U = A \cap \overline{B}$, se tiene que

    $$\begin{aligned} P(A|U) &= \frac{P(A \cap U)}{P(U)} \\ &= \frac{P(A \cap \overline{B})}{P(A \cap \overline{B}) + P(\overline{A} \cap B)} \\ &= \frac{0{,}8 \times 0{,}4}{0{,}8 \times 0{,}4 + 0{,}2 \times 0{,}6} \\ &= 0{,}7272 \end{aligned}$$

• • •

20. Cuatro estudiantes comparten piso, después de una comida deciden el siguiente procedimiento para elegir quién lava los platos: en un sombrero meten cuatro papeles doblados e indistinguibles, en uno de ellos pone 'Sí' y en los otros tres 'No'. El primero coge un papel y lo abre, si el resultado es 'Sí' le toca fregar, si el resultado es que 'No' se libra. No se devuelve el papel al sombrero. A continuacin el segundo coge el papel, y se aplica el mismo criterio, y as sucesivamente hasta que uno de ellos obtiene el papel con el 'Sí'. ¿Quién tiene más probabilidad de lavar los platos, el primero o el último?

    Solución: Llamando N_i al suceso en la extracción i ha salido NO y S_i al suceso ha salido SÍ en la extracción i, los cuatro posibles resultados de este experimento aleatorio son:

$$
\begin{aligned}
P(S_1) &= \frac{1}{4} \\
P(N_1 \cap S_2) &= P(N_1)P(S_2|N_1) \\
&= \frac{3}{4} \times \frac{1}{3} = \frac{1}{4} \\
P(N_1 \cap N_2 \cap S_3) &= P(N_1)P(N_2|N_1)P(S_3|N_1 \cap N_2) \\
&= \frac{3}{4} \times \frac{2}{3} \times \frac{1}{2} = \frac{1}{4} \\
P(N_1 \cap N_2 \cap N_3 \cap S_4) &= P(N_1)P(N_2|N_1)P(N_3|N_1 \cap N_2)P(S_4|N_1 \cap N_2 \cap N_3) \\
&= \frac{3}{4} \times \frac{2}{3} \times \frac{1}{2} \times 1 = \frac{1}{4}
\end{aligned}
$$

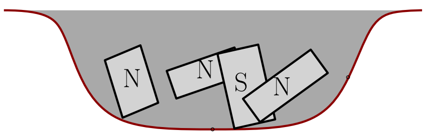


Figura 1.15: Sombrero con cuatro papeletas

Los cuatro estudiantes tienen la misma probabilidad de fregar los platos 1/4. Da igual ser el primero, el segundo, el tercero o el último.

• • •

21. Una comunidad de vecinos ha contratado un sistema de alarma para evitar robos en sus hogares. En caso de robo la alarma se pone en funcionamiento con seguridad. La probabilidad de que se produzca un robo en el vecindario es 0.001. Existen además otras causas (viento, golpes bruscos, ...) que ponen en funcionamiento el sistema de alarma, la probabilidad de que en una noche se produzca una falsa alarma es 0.01. Si se declara una señal de alarma, ¿cuál es la probabilidad de que sea falsa?

    Solución: Llamando R al suceso “se produce un robo”, A al suceso “suena la alarma” y $\overline{R}$ y $\overline{A}$ a sus complementarios, nos piden calcular $P\overline{R}|A)$, teniendo en cuenta que $P(R) = 0{,}001$, $P(A|R) = 1$ y $P(A|\overline{R}) = 0{,}01$. Aplicando el teorema de Bayes:

$$P(\overline{R}|A) = \frac{P(A|\overline{R})P(\overline{R})}{P(A|\overline{R})P(\overline{R}) + P(A|R)P(R)} = \frac{0{,}01 \times 0{,}999}{0{,}01 \times 0{,}999 + 1 \times 0{,}001} = 0{,}9090$$

• • •

22. Una pareja que espera un hijo está preocupada porque un test practicado al feto ha dado positivo a una enfermedad muy rara que solo la padecen uno de cada 10000 individuos. Sin embargo, el test es bastante seguro: de acuerdo con el laboratorio acierta el 99 por ciento de los casos, tanto para los bebés con la enfermedad como para los sanos. ¿Cuál es la probabilidad de que el feto tenga la enfermedad teniendo en cuenta que el resultado del test ha sido positivo?

    Solución: Es el clásico ejercicio de aplicación del teorema de Bayes. Sea S el suceso "el feto está sano" y T el "test ha dado positivo"; sus complementarios son $\overline{S}$, "el feto está enfermo" y $\overline{T}$, "el test ha dado negativo". Las probabilidades conocidas son $P(\overline{S}) = 0{,}0001$, $P(T|\overline{S}) = 0{,}99$ y $P(\overline{T}|S) = 0{,}99$. Se tiene que obtener la $P(S|T)$,

$$P(S|T) = \frac{P(T|S)P(S)}{P(T|S)P(S) + P(T|\overline{S})P(\overline{S})} = \frac{0{,}01 \times 0{,}9999}{0{,}01 \times 0{,}9999 + 0{,}99 \times 0{,}0001} = 0{,}99$$

    La probabilidad de que esté enfermo si el test ha dado positivo es $P(\overline{S}|T) = 1 - P(S|T) = 0{,}01$.

• • •

23. La probabilidad de que un componente se averíe en un período de tiempo dado es 0.01. Su estado (averiado o funcionando) se comprueba con un ensayo que cumple que cuando el componente funciona la probabilidad de que el ensayo diga lo contrario es 0.05, pero si el componente está averiado el ensayo no se equivoca. Si el ensayo indica que el componente está averiado, ¿cuál es la probabilidad de que realmente lo esté?

    Solución: Denominamos F al suceso "la máquina funciona", N al suceso la "máquina no funciona", T_F el "test indica que la máquina funciona" y T_N el "test dice que la máquina no funciona". Los datos que nos proporcionan son $P(N) = 0{,}01$, $P(T_N|F) = 0{,}05$ y $P(T_N|N) = 1$; y nos preguntan por $P(N|T_N)$. Utilizando la fórmula de Bayes:

$$\begin{aligned} P(N|T_N) &= \frac{P(N \cap T_N)}{P(T_N)} \\ &= \frac{P(T_N|N)P(N)}{P(T_N|N)P(N) + P(T_N|F)P(F)} \\ &= \frac{1 \times 0{,}01}{1 \times 0{,}01 + 0{,}05 \times 0{,}99} \\ &= 0{,}168 \end{aligned}$$

    Es decir que la probabilidad de que funcione aunque el test haya indicado avería es muy alta, $P(F|T_N) = 0{,}832$.

• • •

24. Un concursante debe elegir entre tres puertas, detrás de una de las cuales se encuentra un magnífico regalo. Hecha la elección, el presentador que sabe donde se encuentra el premio le abre una de las dos puertas no escogidas donde (lógicamente) no está el premio y a continuación le da al concursante la posibilidad de cambiar la puerta elegida. ¿Qué debe hacer el concursante?

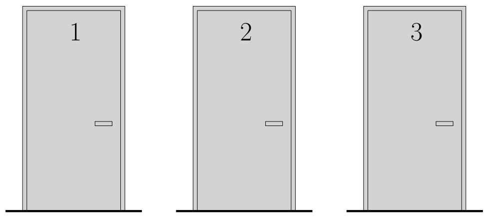


Figura 1.16: El concursante debe elegir una de las tres puertas

Solución: Vamos a suponer que el concursante ha elegido la puerta 1 y el presentador que sabe donde está el regalo, le ha abierto la puerta 2 (ver figura 1.17). La pregunta que queremos contestar es si el concursante debe cambiar de puerta y elegir la 3 o quedarse con la puerta inicial, la número 1.

Se supone que la probabilidad inicial de que el regalo esté detras de cada puerta es la misma e igual a 1/3. Llamando R_i al suceso “el regalo está en la puerta i” y A_i el suceso .el presentador ha abierto la puertas i”, con $i = 1, 2, 3$. Vamos a resolver el caso en el que el concursante elige la puerta 1. El presentador una vez que el concursante ha elegido la puerta le abre una de las otras dos. Lógicamente el presentador nunca abre la puerta que contiene el regalo. Si puede elegir entre dos puertas, lo hace al azar, con probabilidad 50 % para cada puerta.

Voy a calcular la probabilidad de que el regalo esté en la puerta 1, en la 2 y en la 3, si el concursante ha elegido la puerta 1 y el presentador ha abierto la puerta 2. Para simplificar la notación y la explicación no se incluye (aunque se tiene en cuenta) la condición de que el concursante ha elegido la puerta 1.

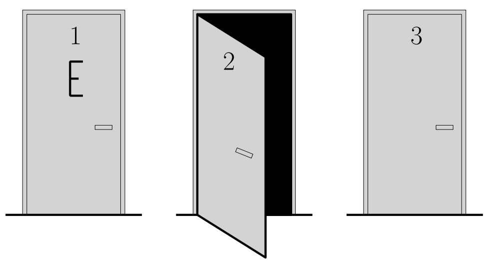


Figura 1.17: El concursante elige la puerta 1 y el presentador abre la puerta 2

La probabilidad de que el regalo esté en la puerta 1 es:

$$
\begin{aligned}
P(R_1|A_2) &= \frac{P(R_1 \cap A_2)}{P(A_2)} \\
&= \frac{P(A_2|R_1)P(R_1)}{P(A_2|R_1)P(R_1) + P(A_2|R_2)P(R_2) + P(A_2|R_3)P(R_3)}
\end{aligned}
$$

Como el concursante ha elegido la puerta 1, $P(A_2|R_1)- = 1/2$ (si el regalo está en la puerta 1, el presentador puede elegir entre la puerta 2 o la 3 al azar), si está en la puerta 2, $P(A_2|R_2) = 0$ (el presentador no puede abrir la puerta si el regalo está en la 2) y si el regalo está en la puerta 3 $P(A_2|R_3) = 1$ (el presentado está obligado a abrir la puerta 2. Ésto es la clave del problema. Además, inicialmente $P(R_1) = P(R_2) = P(R_3) = 1/3$.

$$
P(R_1|A_2) = \frac{\frac{1}{2} \times \frac{1}{3}}{\frac{1}{2} \times \frac{1}{3} + 0 \times \frac{1}{3} + 1 \times \frac{1}{3}} = \frac{1}{3}
$$

De la misma forma, las probabilidades de que el regalo esté en la puerta 2 y 3 es

$$
\begin{aligned}
P(R_2|A_2) &= \frac{P(A_2|R_2)P(R_2)}{P(A_2|R_1)P(R_1) + P(A_2|R_2)P(R_2) + P(A_2|R_3)P(R_3)} \\
&= \frac{0 \times \frac{1}{3}}{\frac{1}{2} \times \frac{1}{3} + 0 \times \frac{1}{3} + 1 \times \frac{1}{3}} = 0
\end{aligned}
$$

$$
\begin{aligned}
P(R_3|A_2) &= \frac{P(A_2|R_3)P(R_3)}{P(A_2|R_1)P(R_1) + P(A_2|R_2)P(R_2) + P(A_2|R_3)P(R_3)} \\
&= \frac{1 \times \frac{1}{3}}{\frac{1}{2} \times \frac{1}{3} + 0 \times \frac{1}{3} + 1 \times \frac{1}{3}} = \frac{2}{3}
\end{aligned}
$$

La probabilidad de que esté en la puerta 3 es el doble, por tanto la recomendación es que el concursante debe cambiar.

• • •

Capítulo 2

Variable aleatoria

1. Dada la variable aleatoria discreta X, cuya función de probabilidad viene definida por

$$P(X = x) = kx, \quad x = 1, 2, ..., 5$$

   a) Calcular el valor de la constante k
   b) Calcular $P(X > 2)$
   c) Calcular $E[X]$ y $Var[X]$.
   d) Calcular $E[Y]$ si $Y = 2X + 5$

Solución:

   a) Como $P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5) = 1$, se tiene que

$$k + 2k + 3k + 4k + 5k = 1$$

$$15k = 1 \quad \Rightarrow \quad k = \frac{1}{15}.$$

   b) $P(X > 2) = \frac{3}{15} + \frac{4}{15} + \frac{5}{15} = \frac{4}{5}$

   c) $E[X] = 1 \times \frac{1}{15} + 2 \times \frac{2}{15} + 3 \times \frac{3}{15} + 4 \times \frac{4}{15} + 5 \times \frac{5}{15} = \frac{11}{3}$

   Para calcular la varianza aplicamos $\mathrm{Var}[X] = E[X^2] - E[X]^2$, donde

$$E[X^2] = 1^2 \times \frac{1}{15} + 2^2 \times \frac{2}{15} + 3^2 \times \frac{3}{15} + 4^2 \times \frac{4}{15} + 5^2 \times \frac{5}{15} = 15$$

   y por tanto

$$\mathrm{Var}[X] = 15 - \left(\frac{11}{3}\right)^2 = \frac{14}{9}$$

d) $E[Y] = 2E[X] + 5 = 2 \times \dfrac{11}{3} + 5 = 12{,}33$

$\mathrm{Var}[Y] = 4\mathrm{Var}[X] = 4 \times \dfrac{14}{9} = 6{,}22$

• • •

2. Dada la variable aleatoria X, cuya función de densidad es

$$f(x) = \begin{cases} k(1 - x^2), & \text{si } 0 < x < 1 \\ 0, & \text{en el resto} \end{cases}$$

a) Obtener k.

b) Dibujar la función de densidad.

c) Calcular la probabilidad $P(X < 0{,}3)$.

d) Obtener la media y la varianza de X.

e) Obtener la media y la varianza de la variable $Y = 3X - 1$.

f) Obtener la media y la varianza de la variable $Z = 3X^2$.

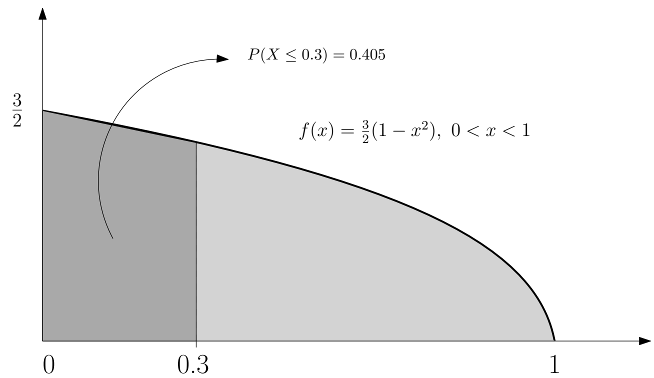


Figura 2.1: Función de densidad del ejercicio 2

Solución:

a) Para que f sea una función de densidad debe cumplir $f(x) \geq 0$ y $\int_{-\infty}^{\infty} f(x)dx = 1$. Según la primera $k > 0$ y según la segunda

$$\int_0^1 k(1 - x^2)dx = 1 \Rightarrow k(x - \frac{x^3}{3})\Big|_0^1 = k(1 - \frac{1}{3}) = 1$$

por tanto $k = 3/2$.

b) Se muestra en la figura 2.1

c) $P(X < 0{,}3) = \displaystyle\int_0^{0{,}3} \frac{3}{2}(1 - x^2)dx = \frac{3}{2}(x - \frac{x^3}{3})\Big|_0^{0{,}3} = 0{,}405$

d) $E[X] = \int_0^1 x \times \frac{3}{2}(1-x^2)dx = \frac{3}{2}(\frac{x^2}{2} - \frac{x^4}{4})\Big|_0^1 = \frac{3}{8}$
$\text{Var}[X] = E[X^2] - E[X]^2,$
$E[X^2] = \int_0^1 x^2 \times \frac{3}{2}(1-x^2)dx = \frac{3}{2}(\frac{x^3}{3} - \frac{x^5}{5})\Big|_0^1 = \frac{2}{10}$
$\text{Var}[X] = \frac{2}{10} - (\frac{3}{8})^2 = 0{,}059$

e) $E[Y] = 3E[X] - 1 = \frac{1}{8}$
$\text{Var}[Y] = 9\text{Var}[X] = 9 \times 0{,}059 = 0{,}534$

f) $E[Z] = 3E[X^2] = \frac{6}{10}$
$\text{Var}[Z] = E[Z^2] - E[Z]^2 = E[9X^4] - (\frac{6}{10})^2$
$E[X^4] = \int_0^1 x^4 \times \frac{3}{2}(1-x^2)dx = \frac{3}{2}(\frac{x^5}{5} - \frac{x^7}{7})\Big|_0^1 = \frac{3}{35}$
$\text{Var}[Z] = 9 \times \frac{3}{35} - (\frac{6}{10})^2 = \frac{72}{175}$

• • •

3. Un gran almacén guarda cajas que contienen piezas de distinto tipo. La proporción X de piezas de tipo A en una caja se puede considerar una variable aleatoria con función de densidad:

$$f(x) = kx(1-x) \text{ con } 0 \leq x \leq 1$$

a) Calcular el valor de k, la media y la varianza de la variable aleatoria X.

b) Si se toman 10 cajas al azar.¿Cuál es la probabilidad de que ninguna de ellas contenga una proporción de piezas de tipo A igual o superior al 75 % ?

Solución:

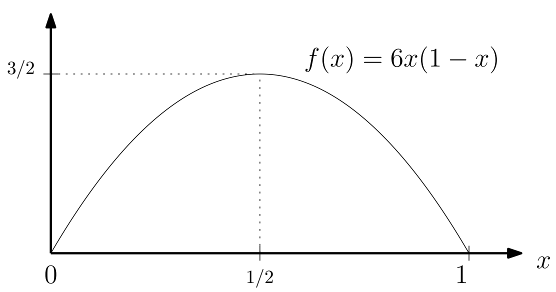


Figura 2.2: Función de densidad de X

a) Para obtener k se pone la condición de que integre 1,

$$\int_0^1 kx(1-x)dx = k\left(\frac{x^2}{2}-\frac{x^3}{3}\right)\Big|_0^1 = k\left(\frac{1}{2}-\frac{1}{3}\right) = \frac{k}{6}$$

como $k/6 = 1$, se tiene que $k = 6$.
La media se obtiene como:

$$E[X] = \int_0^1 x \times \underbrace{6x(1-x)}_{f(x)}dx = 6\left(\frac{x^3}{3}-\frac{x^4}{4}\right)\Big|_0^1 = \frac{1}{2}$$

y la varianza, teniendo en cuenta que $\mathrm{Var}[X] = E[X^2] - E[X]^2$,

$$E[X^2] = \int_0^1 x^2 \times \underbrace{6x(1-x)}_{f(x)}dx = 6\left(\frac{x^4}{4}-\frac{x^5}{5}\right)\Big|_0^1 = \frac{3}{10}$$

$$\mathrm{Var}[X] = \frac{3}{10} - \left(\frac{1}{2}\right)^2 = \frac{1}{20}$$

b) Llamando A_i al suceso "la caja i tiene una proporción de piezas de tipo A menor de 0.75",

$$P(A_i) = \int_0^{0,75} 6x(1-x)dx = 6\left(\frac{x^2}{2}-\frac{x^3}{3}\right)\Big|_0^{0,75} = \frac{27}{32} = 0{,}84375$$

que todas contengan menos de 0.75, significa que

$$P(A_1 \cap A_2 \cap \cdots A_{10}) = P(A_1)P(A_2)\cdots P(A_{10}) = \left(\frac{27}{32}\right)^{10} = 0{,}1828681$$

• • •

4. Un modelo que habitualmente se utiliza en balística para comprobar la correcta calibración de las armas es

$$f(x) = \frac{x}{\sigma^2}\exp\left[-\frac{x^2}{2\sigma^2}\right], \qquad x \geq 0, \sigma \geq 0,$$

donde la variable aleatoria X es la distancia del punto de impacto del proyectil al centro del blanco al que iba dirigido y σ es el parámetro que mide la precisión. Si para una distancia determinada de disparo la precisión del arma es $\sigma = 10$ cm,

a) ¿Cuál es la probabilidad de que un impacto esté a una distancia menor o igual de 5 cm del centro?

b) Calcular la función de distribución.

c) ¿Cuál es la probabilidad de que al lanzar 10 proyectiles, ninguno haya impactado a una distacia menor de 5 cm del centro del blanco?

d) ¿Cuál es la probabilidad de que al lanzar 10 proyectiles, todos hayan impactado a una distacia menor de 5 cm del centro del blanco?

Solución:

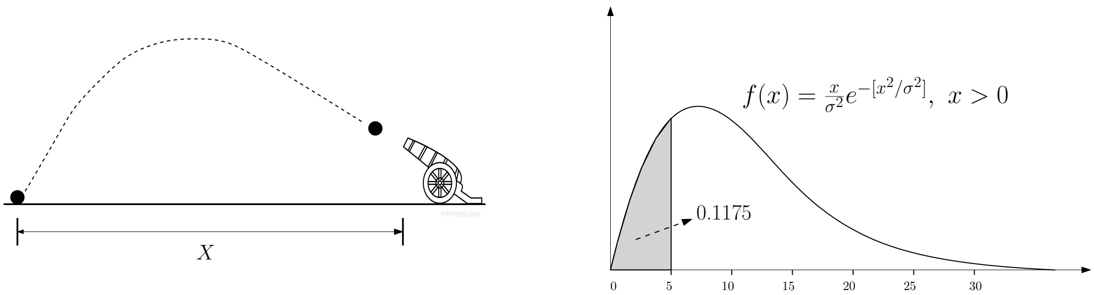


Figura 2.3: Precisión de un cañon

a) $P(X \leq 5) = \int_0^5 \frac{x}{\sigma^2} \exp\left[-\frac{x^2}{2\sigma^2}\right] dx$

haciendo el cambio de variable $u = \frac{x^2}{2\sigma^2}$, $du = \frac{x}{\sigma^2}$ y los nuevos límites de integración son 0 y $u = \frac{5^2}{2 \cdot 100} = \frac{1}{8}$,

$P(X \leq 5) = \int_0^{1/8} e^{-u} du = 1 - e^{-1/8} = 0{,}1175$

Ver figura (2.3).

b)

$$F(x) = P(X \leq x) = \int_0^x \frac{t}{\sigma^2} e^{-t^2/2\sigma^2} dt = -e^{-t^2/2\sigma^2}\Big|_0^x = 1 - e^{-x^2/2\sigma^2}$$

c) Llamamos A_i al suceso "el disparo i está más allá de 5 m del objetivo ". Según hemos visto $P(A_i) = P(X \geq 5) = e^{-25/200} = e^{-1/8}$. La probabilidad de que los diez disparos estén más allá de 5m es

$$P(A_1 \cap A_2 \cap \cdots \cap A_{10}) = P(A_1)P(A_2)\cdots P(A_{10}) = (e^{-1/8})^{10} = e^{-10/8} = 0{,}2865$$

d) La probabilidad de que un disparo se encuentra a una distancia menor o igual de 5 m es $(1 - e^{-1/8})$, y haciendo el mismo razonamiento que en el apartado anterior, la probabilidad de que los diez caigan dentro de 5 metros es $(1 - e^{-1/8})^{10} = 0{,}1175^{10} = 5{,}017567e - 10 \approx 0$.

• • •

5. Una variable aleatoria tiene distribución uniforme en el intervalo $[-a, a]$. Calcular la media, la varianza, el coeficiente de asimetría y el coeficiente de apuntamiento (o curtosis).

Solución: La función de densidad de la variable aleatoria es (ver figura 2.4)

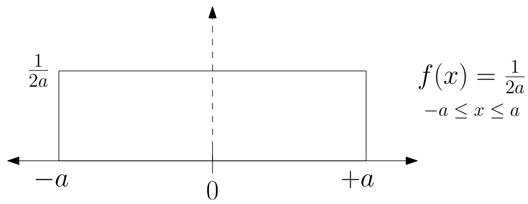


Figura 2.4: Distribución uniforme en $[-a, a]$

$$f(x) = \frac{1}{2a}, -a \leq x \leq a$$

Se observa que el centro de la distribución es 0 (ver figura 2.4).

$$E[X] = \int_{-a}^{a} x \times \frac{1}{2a} dx = \frac{x^2}{4a}\Big|_{-a}^{+a} = 0$$

Para calcular la varianza usamos $\mathrm{Var}[X] = E[X^2] - E[X]^2 = E[X^2]$,

$$E[X^2] = \int_{-a}^{a} x^2 \times \frac{1}{2a} dx = \frac{x^3}{6a}\Big|_{-a}^{+a} = \frac{a^2}{3}$$

por tanto $\mathrm{Var}[X] = a^2/3$. Para obtener el coeficiente de asimetría es necesario calcular $E[(X-\mu)^3]$, como $\mu = E[X] = 0$, tenemos que calcular $E[X^3]$:

$$E[X^3] = \int_{-a}^{a} x^3 \times \frac{1}{2a} dx = \frac{x^4}{8a}\Big|_{-a}^{+a} = 0$$

El coeficiente de asimetría es 0, la distribución es simétrica (ver figura 2.4).

Igual que antes, para calcular el coeficiente de curtosis ahora necesitamos $E[(X-\mu)^4] = E[X^4]$ que es igual a

$$E[X^4] = \int_{-a}^{a} x^4 \times \frac{1}{2a} dx = \frac{x^5}{10a}\Big|_{-a}^{+a} = \frac{a^4}{5}$$

El coeficiente de curtosis o apuntamiento es

$$\kappa = E[(X-\mu)^4]/(E[(X-\mu)^2])^2 = (a^4/5)/(a^2/3)^2 = \frac{9}{5}$$

El coeficiente de curtosis es siempre mayor que 1, en este caso es $\kappa = 1{,}8$. La distribución normal se utiliza de referencia, tiene coeficiente de apuntamiento $\kappa = 3$. La distribución con menor coeficiente de apuntamiento es

$$P(X = x) = \begin{cases} 1/2, & x = -1 \\ 1/2, & x = +1 \end{cases}$$

que es una distribución simétrica de media cero que cumple $E[X^k] = 0$ si k es impar y $E[X^k] = 1$ si k es par, y por tanto $\kappa = E[X^4]/(E[X^2])^2 = 1$.

• • •

6. Si X es una variable aleatoria con media μ. Demostrar que cuando $m = \mu$, $E[(X - m)^2]$ es mínima.

   Solución: Vamos a resolverlo de dos formas:

   La primera forma es calculando el mínimo de $g(m) = E[(X - m)^2]$

$$\begin{aligned} g(m) &= E[X^2 + m^2 - 2mX] \\ &= E[X^2] + m^2 - 2mE[X] \\ &= (\sigma^2 + \mu^2) + m^2 - 2\mu m \end{aligned}$$

   llamando μ y σ^2 a la media y la varianza de X. Hemos utilizado la fórmula $\text{Var}[X] = E[X^2] - E[X]^2$.

   Tenemos que $g(m) = (\sigma^2 + \mu^2) + m^2 - 2\mu m$, derivando e igualando a cero

$$\frac{dg(m)}{dm} = 2m - 2\mu \quad \Rightarrow \quad 2m - 2\mu \quad \Rightarrow \quad m = \mu$$

   vamos a comprobar que es un mínimo:

$$\frac{d^2g(m)}{dm^2} = 2 > 0 \Rightarrow \text{en } m = \mu, \text{ por tanto hay un mínimo}$$

   Otra forma

$$\begin{aligned} E[(X - m)^2] &= E[(X - \mu + \mu + m)^2] \\ &= E[(X - \mu)^2 + (\mu - m)^2 + 2(\mu - m)(X - \mu)] \\ &= \underbrace{E[(X - \mu)^2]}_{\sigma^2} + \underbrace{E[\mu - m)^2]}_{(\mu - m)^2} + 2(\mu - m)\underbrace{E[(X - \mu)]}_{=0} \\ &= \sigma^2 + (\mu - m)^2 \end{aligned}$$

   Como $(\mu - m)^2 \geq 0$, el valor anterior es mínimo para $\mu = m$.

• • •

7. Demostrar la desigualdad de Markov

$$P(X \geq a) \leq \frac{1}{a} E\left[X\right]$$

donde X es una variable aleatoria positiva ($P(X > 0) = 1$).

Solución: Esta desigualdad en algunos libros de texto se denomina también de Chebychev. Es importante e imprescindible la condición de que los valores que toma la v.a. X son positivos.

$$E[X] = \int_0^\infty x f_X(x) dx = \underbrace{\int_0^a x f_X(x) dx}_{(1)} + \underbrace{\int_a^\infty x f_X(x) dx}_{(2)}$$

tanto (1) como (2) son positivos, por tanto si quitamos (1) tenemos

$$E[X] \geq \int_a^\infty x f_X(x) dx.$$

En la integra $x \in [a, \infty)$, por tanto $x \geq a$ y

$$\int_a^\infty x f_X(x) dx \geq a \int_a^\infty f_X(x) dx$$

Teniendo en cuenta que $P(X \geq a) = \int_a^\infty f_X(x) dx$, se tiene

$$E[X] \geq a P(X \geq a),$$

$$P(X \geq a) \leq \frac{E[X]}{a} \tag{2.1}$$

Que es lo que queríamos demostrar. Vamos a aadir algo más: si llamamos $\mu = E[X]$ y $a = k\mu$, la desigualdad 2.1 se puede poner como

$$P(X \geq k\mu) \leq \frac{1}{k}.$$

A partir de lo anterior se puede demostrar la desigualdad más conocida de Chebychev. Sea X una variable aleatoria con media μ y varianza σ^2 (ahora no tiene que ser positiva) se define $Y = (X - \mu)^2$ y aplicamos el resultado anterior 2.1, con $a = k^2$ se tiene

$$P((X - \mu)^2 \geq k^2 \sigma^2) \leq \frac{1}{k^2},$$

que se puede escribir como

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}.$$

Esta desigualdad de Chebychev es fundamental para demostrar algunos teoremas de convergencia en probabilidad.

• • •

8. De acuerdo con la teoría cinética de los gases, la velocidad V de una molécula de masa m de un gas a la temperatura (absoluta) T es una variable aleatoria con la siguiente función de densidad:

$$f(v) = \frac{4}{\alpha^3\sqrt{\pi}} v^2 e^{-v^2/\alpha^2}, \quad v \geq 0$$

donde $\alpha = \sqrt{2kT/m}$, siendo k la constante de Boltzmann. Además,

$$E(V) = 2\alpha/\sqrt{\pi}, \quad \mathrm{Var}(V) = (3/2 - 4/\pi)\alpha^2$$

a) Calcular el valor medio de la energía cinética, $mV^2/2$, de una molécula. ¿ A una misma temperatura T, qué gas tiene mayor valor medio de energía cinética, uno ligero u otro más pesado?

b) Obtener la función de densidad de la energía cinética de una molécula. Indicar si depende de la masa molecular.

Solución:

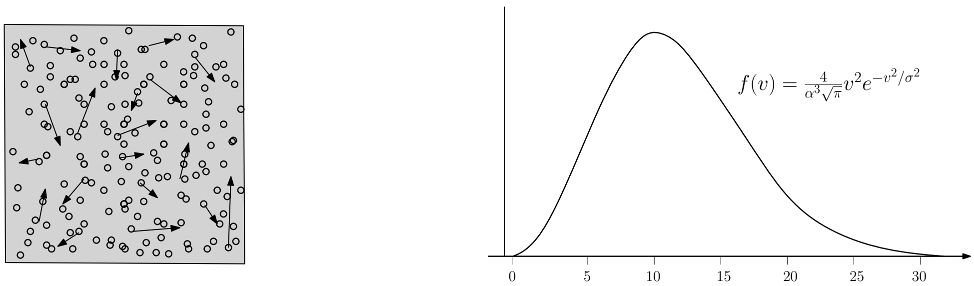


Figura 2.5: Distribución de Maxwell

a) Si $Y = \frac{1}{2}mV^2$, su esperanza es $E[Y] = \frac{1}{2}mE[V^2]$. Teniendo en cuenta que $\mathrm{Var}[V] = E[V^2] - E[V]^2$,

$$\begin{aligned} E[V^2] &= \mathrm{Var}[V] + E[V]^2 \\ &= (3/2 - 4/\pi)\alpha^2 + (2\alpha/\sqrt{\pi})^2 \\ &= \frac{3}{2}\alpha^2 \\ &= \frac{3kT}{m} \\ E[Y] &= \frac{1}{2}mE[V^2] = \frac{1}{2}m \times \frac{3kT}{m} = \frac{3}{2}kT \end{aligned}$$

Según lo anterior, la energía cinética media de un gas ideal no depende de la masa, solo depende de la temperatura. Dos gases diferentes a la misma temperatura tienen la misma energía cinética (media).

b)

$$
\begin{aligned}
F_Y(y) &= P(Y \leq y) = P(\frac{1}{2}mV^2 \leq y) = P(V \leq \sqrt{\frac{2y}{m}}) \\
f_Y(y) &= \frac{dF_Y(y)}{dy} = \frac{1}{2}\sqrt{\frac{2}{m}}\frac{1}{\sqrt{y}}f(\sqrt{\frac{2y}{m}}) \\
&= \frac{4}{m^{3/2}\alpha^3}\sqrt{\frac{2y}{\pi}}e^{-\frac{y}{kT}}
\end{aligned}
$$

sustituyendo $\alpha = \sqrt{2kT/m}$

$$
f_y(y) = 2\left(\frac{1}{kT}\right)^{3/2}\sqrt{\frac{y}{\pi}}e^{-\frac{y}{kT}}
$$

La distribución de probabilidad de Y (energía cinética) solo depende de la temperatura T y no depende de la masa del gas ideal. Los gases más pesados se mueven lentamente y los más ligeros con mayor rapidez.

• • •

9. Sea X una variable aleatoria con distribución uniforme en $(0, 1)$. Calcular la probabilidad de que $Y > 0{,}8$ si $Y = e^{-X^2}$.

Solución: $P(Y > 0{,}8) = P(e^{-X^2} > 0{,}8) = P(X < \sqrt{-\log 0{,}8})$ La variable aleatoria X tiene distribución de probabilidad uniforme en el intervalo (0,1). Por tanto cumple que $P(X \leq x) = x$ por tanto,

$$
P(Y > 0{,}8) = P(X < \sqrt{-\log 0{,}8}) = -\log 0{,}8 = 0{,}4724
$$

La clave de los ejercicios de cambio de variable es despejar X, en este caso, en la desigualdad $e^{-X^2} > 0{,}8$ (ver figura 2.6)

$$
\begin{aligned}
&e^{-X^2} > 0{,}8 \\
&-X^2 > \log 0{,}8 \quad \Rightarrow \quad X^2 < -\log 0{,}8 \\
&X < \sqrt{-\log 0{,}8}
\end{aligned}
$$

• • •

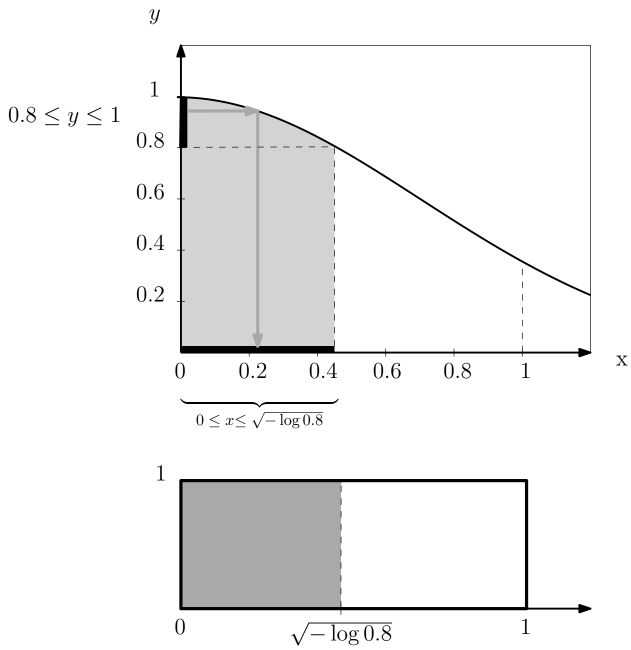


Figura 2.6: Probabilidad con cambio de variable

10. La función de distribución de la variable aleatoria X es $F_X(x)$. Obtener la función de densidad de la variable aleatoria $Y = F_X(X)$.

    Solución: Este resultado es útil en varias aplicaciones. La variable aleatoria X tiene función de distribución F_X y definimos una nueva variable aleatoria con esta misma función $Y = F_X(X)$ (la notación es importante, X e Y en mayúsculas). Vamos a obtener la función de distribución de Y:

    $$F_Y(y) = P(Y \leq y) = P(F_X(X) \leq y) = P(X \leq F_X^{-1}(y))$$

    La solución de $F_X(X) \leq y$ es $X \leq F_X^{-1}(y)$ pues F_X es creciente por ser una función de distribución:

    $$F_Y(y) = P(X \leq F_X^{-1}(y)) = F_X(F_X^{-1}(y)) = y, \quad 0 \leq y \leq 1$$

    Los límites $0 \leq y \leq 1$ pues $0 \leq F_X(x) \leq 1$. La función de densidad de Y es

    $$f_Y(y) = \frac{dF_Y(y)}{dy} = 1, \quad 0 \leq y \leq 1$$

    .

    Y tiene distribución uniforme en $[0, 1]$.

• • •

11. La variable aleatoria Z tiene como función de densidad

$$f_Z(z) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{z^2}{2\sigma^2}\right), \quad -\infty < z < \infty, \quad \sigma > 0$$

Obtener la función de densidad de $Y = |Z|$ y su media.

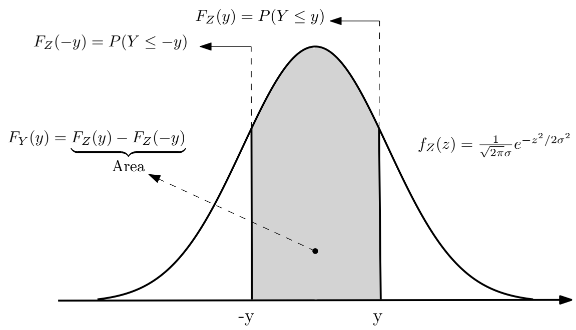


Figura 2.7: Distribución valor absoluto de una v.a. normal de media 0

Solución: Empezamos obteniendo la función de distribución de Y

$$\begin{aligned}
F_Y(y) &= P(Y \leq y) && \text{(definición de } F_Y\text{)} \\
&= P(|Z| \leq y) && \text{(sustituimos } Y \text{ por } |Z|\text{)} \\
&= P(-y \leq Z \leq y) && \text{(despejamos Z)} \\
&= P(Z \leq y) - P(Z \leq -y) && \text{(ver figura 2.7)} \\
&= F_Z(y) - F_Z(-y)
\end{aligned}$$

derivamos la función de distribución

$$\begin{aligned}
f_Y(y) &= \frac{dF_Y(y)}{dy} \\
&= \frac{dF_Z(y)}{dy} - \frac{dF_Z(-y)}{dy} \\
&= f_Z(y) + f_Z(-y) \\
&= \frac{1}{\sqrt{2\pi}\sigma} e^{-y^2/2\sigma^2} + \frac{1}{\sqrt{2\pi}\sigma} e^{-(-y)^2/2\sigma^2} \\
&= \frac{2}{\sqrt{2\pi}\sigma} e^{-y^2/2\sigma^2}, \quad y \geq 0
\end{aligned}$$

Y el valor esperado de Y es

$$E[Y] = \int_0^\infty y \times \frac{2}{\sqrt{2\pi}\sigma} e^{-y^2/2\sigma^2} dy$$

haciendo el cambio de variable $u = \dfrac{y^2}{2\sigma^2}$, con $du = \dfrac{y}{\sigma^2}dy$

$$E[Y] = \sqrt{\frac{2}{\pi}}\sigma \int_0^{\infty} e^{-u} du = \sqrt{\frac{2}{\pi}}\sigma.$$

• • •

12. Si X es una variable aleatoria con distribución uniforme entre 0 y θ, obtener la función de densidad de la variable aleatoria $Y = +\sqrt{X}$.

Solución:

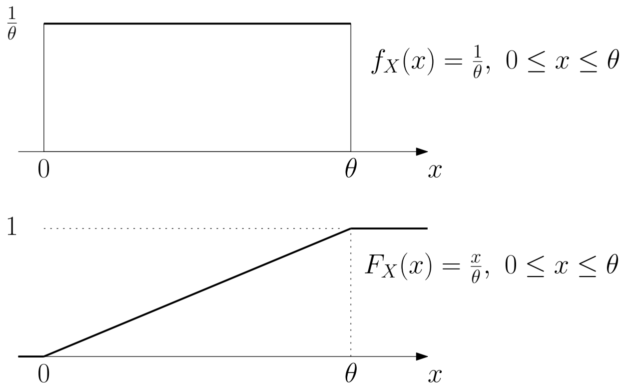


Figura 2.8: Distribución uniforme entre 0 y θ

$$\begin{aligned} F_Y(y) &= P(Y \leq y) \\ &= P(\sqrt{X} \leq y) \\ &= P(X \leq y^2) \\ &= F_X(y^2) \end{aligned}$$

Como X es uniforme, su función de distribución es (ver figura 2.8) $F_X(x) = P(X \leq x) = \dfrac{x}{\theta}$, $0 \leq x \leq \theta$, por tanto

$$F_Y(y) = F_X(y^2) = \frac{y^2}{\theta}, \quad 0 \leq y \leq \sqrt{\theta}$$

y derivando se obtiene la función de densidad de Y,

$$f_Y(y) = \frac{2y}{\theta}, \quad 0 \leq y \leq \sqrt{\theta}.$$

• • •

13. Se elige un punto al azar interior a la circunferencia de ecuación $x^2 + y^2 = r^2$. Llamando Z a la variable aleatoria definida por la distancia entre el punto elegido y el centro de la circunferencia, calcular las funciones de densidad y distribución de Z.

    Solución: Todos los puntos que cumplen $Z \leq z$, se encuentran en el círculo de radio z (ver figura 2.9). Como todos los puntos tienen la misma probabilidad, se tiene que

    $$F_Z(z) = P(Z \leq z) = \frac{\pi z^2}{\pi r^2} = \left(\frac{z}{r}\right)^2, \quad 0 \leq z \leq r$$

    y derivando respecto de z se obtiene la función de densidad

    $$f_Z(z) = \frac{2z}{r^2}, 0 \leq z \leq r.$$

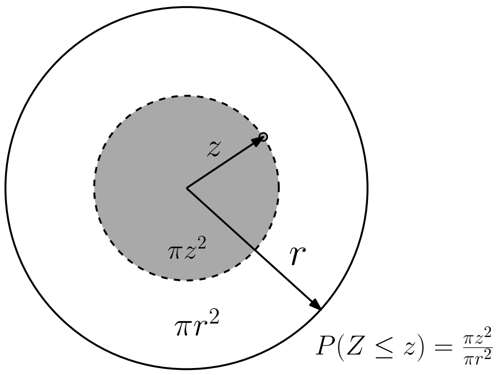


Figura 2.9: Probabilidad de $Z \leq z$

• • •

14. Supóngase una diana circular con centro en el origen de coordenadas de radio r y X, Y las coordenadas de un punto elegido al azar (por ejemplo, el lanzamiento de un dardo). Supóngase que cualquier otro punto de la diana tiene la misma probabilidad de ser elegido. Calcule $f_{XY}(x, y)$, las funciones de densidad marginales y condicionadas.

    Solución: Si cualquier punto interior al círculo $x^2 + y^2 \leq r^2$ tiene la misma probabilidad, significa que la función de densidad conjunta de X e Y es constante k en ese círculo, $f_{X,Y}(x, y) = k$. El volumen comprendido por la función de densidad debe ser 1, por tanto volumen $= \pi r^2 \times k = 1$, implica que $k = 1/\pi r^2$.

    La función de densidad conjunta es

    $$f_{X,Y}(x, y) = \frac{1}{\pi r^2}, \quad x^2 + y^2 \leq r^2$$

    La función de densidad marginal de X es (ver figura 2.11):

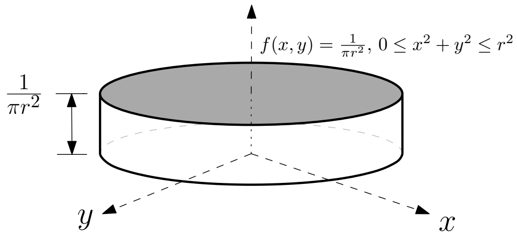


Figura 2.10: Distribución uniforme en un círculo

$$\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy \\
&= \int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} \frac{1}{\pi r^2} dy \\
&= \left.\frac{y}{\pi r^2}\right|_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} \\
&= \frac{2}{\pi r^2}\sqrt{r^2-x^2}, \quad -r \leq x < r
\end{aligned}$$

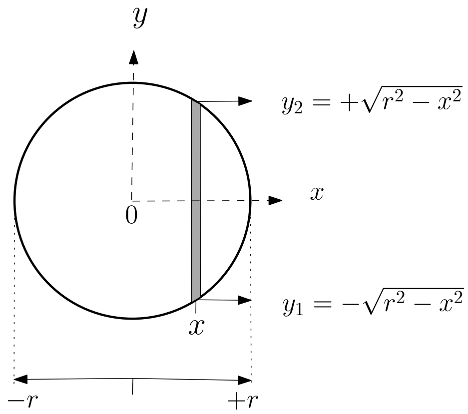


Figura 2.11: Distribución uniforme en un círculo

La función de densidad marginal de Y se obtiene de la misma manera y es igual a:

$$f_Y(y) = \frac{2}{\pi r^2}\sqrt{r^2-y^2}, \quad -r \leq y < r.$$

La función de densidad de X condicionadas a $Y = y$ es:

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{1}{2\sqrt{r^2-y^2}}, \quad -\sqrt{r^2-y^2} \leq x \leq \sqrt{r^2-y^2}$$

En la función condicionda $f_{X|Y}(x|y)$ la variable principal es x, la magnitud y es un valor conocido y constante. Teniendo en cuenta lo anterior, es fácil comprobar que la distribución de X condicionada a $Y = y$ es constante en el intervalo $-\sqrt{r^2 - y^2} \leq x \leq \sqrt{r^2 - y^2}$. Por ejemplo para los valores de $r = 1$ e $y = 0{,}8$ se tiene que

$$f_{X|Y}(x|y = 0{,}8) = \frac{1}{1{,}2}, \quad -0{,}6 \leq x \leq 0{,}6$$

Se observa que la función anterior es una función de densidad y cumple que es no-negativa y que el área bajo la función y el eje x es 1.

El problema es simétrico para Y, la función de densidad de Y condicionadas a $X = x$ es:

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{1}{2\sqrt{r^2 - x^2}}, \quad -\sqrt{r^2 - x^2} \leq y \leq \sqrt{r^2 - x^2}$$

ahora la variable es y y x es un valor constante conocido.

• • •

15. Obtén la distribución de probabilidad del máximo, del mínimo y de la suma de los resultados obtenidos al lanzar dos dados equilibrados. Se acepta que los resultados de los dados son variables aleatorias independientes.

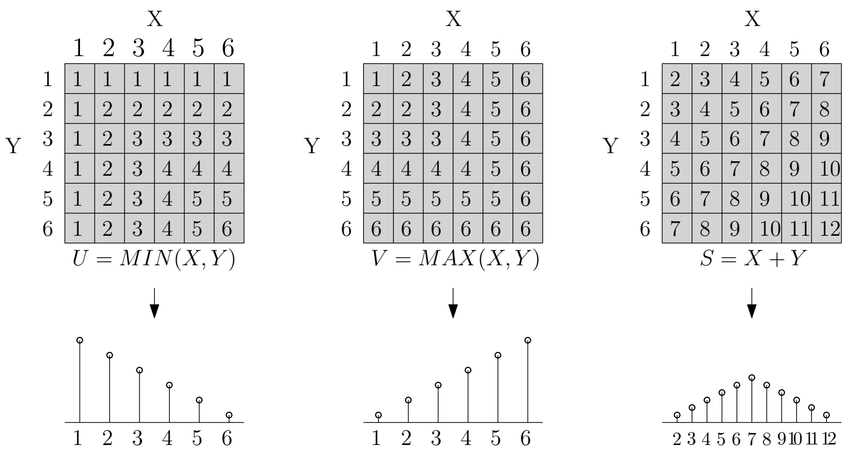


Figura 2.12: Distribución del máximo, mínimo y suma de dos dados

Solución: En este ejercicio se aborda un problema muy importante en probabilidad: la distribución del máximo y del mínimo de valores obtenidos (extraídos) al azar de una distribibión de probabilidad. Cuando la variable aleatoria es discreta y el número de sus posibles resultados es pequeño, se puede resolver el problema por enumeración. En la tablas de la figura 2.12 se muestran los resultados del mínimo (U), máximo (V) y la suma (S) al lanzar dos dados representados por las variables aleatorias X e Y. Cada tabla se muestran los seis resultados de X y los seis de Y, las 36 combinaciones son equiprobables. La tabla de la izquierda (figura 2.12) nos muestra que de las 36 combinaciones, hay 11 que dan lugar a un mínimo igual a 1.

Por tanto, al lanzar los dados, la probabilidad de que el mínimo sea 1 es 11/36. La probabilidad de que el mínimo sea 2 es 9/36, pues de las 36 combinaciones hay 9 que tienen mínimo igual a 2. Así sucesivamente. En la figura de abajo se muestran gráficamente la distribución del mínimo (U).

De la misma manera se pueden obtener la distribución del máximo de los resultados (V) que se muestra en la tabla y gráfico del centro y de la suma que se muestra en la tabla y gráfico de la derecha.

Vamos a obtener las distribuciones de probabilidad de forma analítica. El método que vamos a aplicar se puede utilizar a cualquier pareja de variables aleatorias independientes, ya sean continuas o discretas. Empezamos con la distribución de máximo que es más sencilla. Nos basamos en la siguiente propiedad: si X e Y son variables aleatorias independientes y $V = \text{máx}\,(X, Y)$ se cumple que

$$P(V \leq v) = P(X \leq v, Y \leq v) = P(X \leq v)P(Y \leq v)$$

En palabras, lo que viene a decir la propiedad anterior, es que si el máximo de dos números es "mayor o igual que v" los dos números son mayores o iguales que v. Teniendo en cuenta que $P(X \leq v) = P(Y \leq v) = v/6$ se tiene que

$$P(V \leq v) = \left(\frac{v}{6}\right)^2$$

Este resultado se puede comprobar en la tabla del centro de la figura 2.12. Para calcular la distribución puntual de probabilidad $P(V = v)$ en una variable discreta a partir de la función de distribución se puede aplicar la siguiente relación:

$$P(V = v) = P(V \leq v) - P(V \leq v - 1) = \left(\frac{v}{6}\right)^2 - \left(\frac{v-1}{6}\right)^2 = \frac{2v-1}{36}$$

La distribución del mínimo se obtiene de forma similar, pero utilizando la siguiente variante de la propiedad anterior:

$$P(U \geq u) = P(X \geq u, Y \geq u) = P(X \geq u)P(Y \geq u)$$

Para calcular $P(X \geq u)$ utilizamos

$$P(X \geq u) = 1 - P(X \geq u - 1) = 1 - \frac{u-1}{6} = \frac{7-u}{6},$$

por tanto

$$P(U \geq u) = \frac{(7-u)^2}{36}, \quad u = 1, 2, 3, 4, 5, 6$$

La función de distribución de U será

$$P(U \leq u) = 1 - P(U \geq u - 1) = 1 - \frac{(6-u)^2}{36}$$

La distribución de probabilidad se obtiene teniendo en cuenta que

$$\begin{aligned} P(U=u) &= P(U \le u) - P(U \le u-1) \\ &= \frac{(7-u)^2}{36} - \frac{(6-u)^2}{36} \\ &= \frac{13-2u}{36} \qquad u = 1,2,3,4,5,6 \end{aligned}$$

Como se ha indicado previamente, en problemas de tamaño reducido lo mejor es completar la tabla y obtener a partir de ahí la distribución de la suma. En general, la distribución de la suma de dos variables aleatorias discretas e independientes se puede obtener utilizando

$$P(S=k) = P(X+Y=k) = \sum_{-\infty}^{\infty} P(X=i, Y=k-i) = \sum_{-\infty}^{\infty} P(X=i)P(Y=k-i),$$

en el caso de los dados

$$P(S=k) = \sum_{i=\text{máx}(1,k-6)}^{\text{mín}(6,k-1)} P(X=i)P(Y=k-i), \quad k = 2,3,...,12$$

$$P(S=k) = \begin{cases} \dfrac{k-1}{36}, & 2 \le k \le 7 \\ \dfrac{13-k}{36}, & 8 \le k \le 12 \end{cases}$$

• • •

16. X e Y son dos variables aleatorias continuas independientes con la misma función de distribución F. Calcular la función de densidad de

$$U = \max(\text{X}, \text{Y}).$$

Aplicarlo al caso en que las dos variables tengan distribución uniforme entre [0,1].

Solución: Como se ha visto en el ejercicio 15, aplicamos la siguiente propiedad

$$F_U(u) = P(U \le u) = P(X \le u)P(Y \le u) = F(u) \times F(u) = F(u)^2$$

Si se desea obtener la función de densidad

$$f_U(u) = \frac{dF_U(u)}{du} = 2f(u)F(u)$$

siendo f la función de densidad de X (e Y).

Vamos a aplicarlo al caso de la distribución uniforme, donde $F(x) = x,\ 0 \leq x \leq 1$

$$F_U(u) = F(u)^2 = u^2, 0 \leq u \leq 1$$

y la función de densidad

$$f_U(u) = \frac{dF_U(u)}{du} = 2u, 0 \leq u \leq 1.$$

Se puede extender el resultado a n variables aleatorias $X_1, X_2, ..., X_n$ independientes con la misma distribución de probabilidad F, $U = \text{máx}(X_1, X_2, ..., X_n)$

$$\begin{aligned} F_U(u) &= P(U \leq u) \\ &= P(X_1 \leq u)P(X2 \leq u) \cdots P(X_n < u) \\ &= F(u) \times F(u) \times \cdots \times F(u) \\ &= F(u)^n \end{aligned}$$

La función de densidad será

$$f_U(u) = nf(u)F(u)^{n-1}$$

en el caso de que las variables aleatorias tengan distribución uniforme en el intervalo $[0, 1]$, tenemos que

$$F_U(u) = u^n, \quad \text{y} \quad f_U(u) = nu^{n-1}, 0 \leq u \leq 1.$$

• • •

17. $X_1, X_2, ...X_n$ son n variables aleatorias continuas independientes con la misma función de distribución F. Calcular la función de densidad de

$$V = \min(\mathrm{X_1, X_2, ..., X_n}).$$

Aplicarlo al caso en que las variables tengan distribución $f(x) = \lambda e^{-\lambda x}$ con $\lambda > 0$ y $x \geq 0$.

Solución: Utilizamos la propiedad vista en el ejercicio 15,

$$\begin{aligned} P(V \geq v) &= P(X_1 \geq v, X_2 \geq v, ..., X_n \geq v) \\ &= P(X_1 \geq v)P(X_2 \geq v) \cdots P(X_n \geq v) \\ &= (1 - F(u))^n \end{aligned}$$

por tanto

$$F_V(v) = P(V \leq v) = 1 - (1 - F_X(v))^n,$$

aplicado al caso en el que X_k tiene distribución exponencial $F_X(x) = 1 - e^{-\lambda x}$, dada

$$F_V(v) = 1 - e^{-n\lambda v}, v \geq 0$$

Y la función de densidad es

$$f_V(v) = f(v)\,(1 - F_X(v))^{n-1}$$

en el caso de la distribución exponencial $f_V(v) = n\lambda e^{-n\lambda v}, v \geq 0$.

• • •

18. La función de densidad de una variable aleatoria bidimensional viene dada por la expresión:

$$f_{XY}(x,y) = \begin{cases} xy + ce^x, & \text{cuando } 0 < x < 1 \text{ y } 0 < y < 1 \\ 0, & \text{en el resto} \end{cases}$$

Indicar si son independientes las variables aleatorias X e Y.

Solución: Vamos a obtener las distribuciones marginales f_X y f_Y,

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)dy = \int_0^1 (xy + ce^x)dy = \frac{1}{2}xy^2 + cye^x|_0^1 = \frac{1}{2}x + c\,e^x, 0 \leq x \leq 1.$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)dx = \int_0^1 (xy + ce^x)dx = \frac{1}{2}x^2y + ce^x|_0^1 = \frac{1}{2}y + c(e-1), 0 \leq y \leq 1.$$

Es fácilmente verificable que $f_{XY}(x,y) \neq f_X(x)f_Y(y)$, por tanto no son independientes. Lo anterior no depende del valor de la constante c. El valor de c se puede obtener con la condición

$$\int_0^1 \left(\frac{1}{2}y + c(e-1)\right) dy = \frac{1}{4}y^2 + c(e-1)y|_0^1 = \frac{1}{4} + c(e-1) = 1$$

que proporciona a c el valor $\dfrac{3}{4(e-1).}$

• • •

19. La cantidad en miligramos de dos componentes contenidos en un producto es una variable aleatoria bidemensional, cuya función de densidad viene dada por la expresión

$$f_{XY}(x,y) = \begin{cases} 4xy, & \text{cuando } 0 \leq x \leq 1 \text{ y } 0 \leq y \leq 1 \\ 0, & \text{en el resto} \end{cases}$$

a) Obtener las funciones de densidad marginales

b) Calcular la covarianza

c) Calcular $P(X + Y \leq 0{,}5)$

d) Calcular la probabilidad de que la cantidad del primer componentes sea menor que 0.3 miligramos cuando la del segundo es 0.8 miligramos.

Solución:

a) Las marginales se obtienen de la siguiente forma

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)dy = \int_0^1 4xydy = 2x, 0 \leq x \leq 1.$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)dx = \int_0^1 4xydx = 2y, 0 \leq y \leq 1.$$

b) Se observa que $f_{XY}(x,y) = f_X(x)f_Y(y)$, por tanto X e Y son variables aleatorias independientes, es decir $\mathrm{Cov}(X,Y) = 0$.

c) Para obtener esta probabilidad es recomendable utilizar un gráfico para establecer la región de integración. En la figura 2.13 se determina la región $x \geq 0, y \geq 0, x + y \leq 0{,}5$, tenemos que obtener el volumen encerrado por la función de densidad conjunta en esa región.

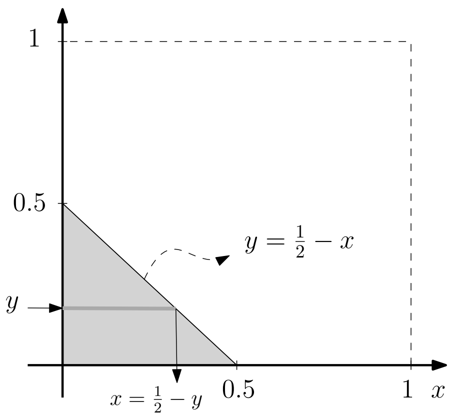


Figura 2.13: Función de densidad de X

$$P(X+Y \leq \frac{1}{2}) = \int_0^{1/2} \left(\int_0^{\frac{1}{2}-y} 4xydx \right) dy \tag{2.2}$$

La integral interior es

$$\int_0^{1/2-y} 4xydx = 2yx^2\Big|_{x=0}^{x=\frac{1}{2}-y} = 2y(\frac{1}{2}-y)^2 = 2y^3 - 2y^2 + \frac{y}{2}$$

Sustituyendo en (2.2) se tiene

$$P(X+Y \leq \frac{1}{2}) = \int_0^{1/2} \left(2y^3 - 2y^2 + \frac{y}{2} \right) dy = 2\frac{y^4}{4} - 2\frac{y^3}{3} + \frac{y^2}{4}\Big|_0^{\frac{1}{2}} = \frac{1}{96}$$

d) Como son independientes

$$P(X \leq 0{,}3|Y = 0{,}8) = P(X \leq 0{,}3) = \int_0^{0{,}3} 2xdx = x^2\Big|_0^{0{,}3} = 0{,}09$$

• • •

20. La función de densidad de la variable aleatoria bidimensional (X, Y), bien dada por la expresión:

$$f_{XY}(x,y) = \begin{cases} kxy, & \text{cuando } 0 < x < y < 1 \\ 0, & \text{en el resto} \end{cases}$$

a) Calcular el valor de k y las distribuciones marginales.

b) Calcular $P(X < 0{,}5|Y = 0{,}5)$.

c) ¿Son independientes las variables aleatorias X e Y?

Solución: En la figura 2.14 se ha representado la región con probabilidad distinto de cero.

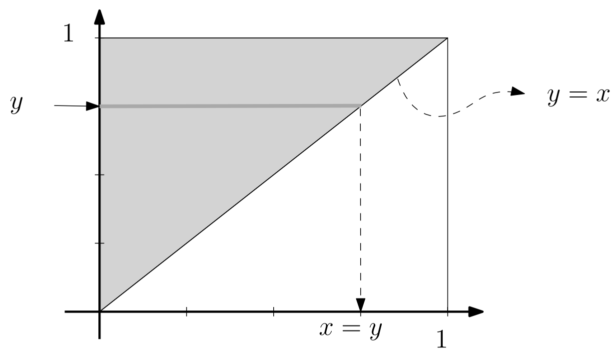


Figura 2.14: Región con probabilidad no nula (ejercicio 20)

.

a) Vamos a calcular las distribuciones marginales y a partir de ellas el valor de k.

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)dy = \int_x^1 kxydy = kx\frac{y^2}{2}\Big|_x^1 = \frac{k}{2}x(1-x^2), \quad 0 \leq x \leq 1.$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)dx = \int_0^y kxydx = ky\frac{x^2}{2}\Big|_0^y = \frac{k}{2}y^3, \quad 0 \leq y \leq 1.$$

Podemos elegir cualquiera de las dos marginales para determinar k. Tenemos que

$$\int_{-\infty}^{\infty} f_Y(y)dy = 1 \Rightarrow \int_0^1 \frac{k}{2}y^3 dy = \frac{k}{2}\frac{y^4}{4}\Big|_0^1 = \frac{k}{8}$$

por lo tanto $k = 8$. Se puede comprobar en la marginal de X,

$$\int_0^1 4x(1-x^2)dx = 4(\frac{x^2}{2} - \frac{x^4}{4})\Big|_0^1 = 1.$$

b) Primero obtenemos la distribución condicionada

$$f_{X|Y}(x|y) = 8xy/4y^3 = 2x/y^2, \; 0 \leq x \leq y.$$

Muy importante tener en cuenta que es una función de densidad de x, el valor y es conocido. Para $y = 0{,}5$,

$$f_{X|Y}(x|y = 0{,}5) = 8x, \quad 0 \leq x \leq 0{,}5$$

La probabilidad que nos piden es 1, no hace falta hacer la integral,

$$P(X < 0{,}5|Y = 0{,}5) = \int_0^{0{,}5} 8xdx = 4x^2\Big|_0^{0{,}5} = 1.$$

c) No son independientes. Basta con observar la región de probabilidad no nula en la figura 2.14 para darse cuenta que la probabilidad de Y depende de X. Lógicamente el producto de las marginales es diferente a la distribución conjunta.

• • •

21. X e Y son variables aleatorias con coeficiente de correlación lineal $\rho = -1$. Si las varianzas son iguales, calcular la varianza de $Z = X + Y - 1$.

Solución: Tenemos en cuenta las siguientes propiedades

a) $\mathrm{Var}(aX + b) = a^2 \,\mathrm{Var}(X)$

b) $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X, Y)$

Llamando $U = X + Y$, tenemos que $\mathrm{Var}(U-1) = \mathrm{Var}(U)$ por la propiedad (a). El coeficiente de correlación de dos variables aleatorias es por definición

$$\rho = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}\,(X)\,\mathrm{Var}\,(Y)}}$$

por tanto, si $\rho = -1$, $\mathrm{Cov}(X,Y) = -\sqrt{\mathrm{Var}\,(X)\,\mathrm{Var}\,(Y)}$ y como las varianzas son iguales $\mathrm{Cov}(X,Y) = -\mathrm{Var}\,(X)$.

En vista de todo ello,

$$\begin{aligned}
\mathrm{Var}(Z) &= \mathrm{Var}(U) \\
&= \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X,Y) \\
&= \mathrm{Var}(X) + \mathrm{Var}(X) - 2\,\mathrm{Var}(X) = 0
\end{aligned}$$

• • •

22. La distribución de probabilidad conjunta de las variables aleatorias Y_1 e Y_2 es la siguiente:

		Y_1		
		-1	0	1
	-1	1/16	3/16	1/16
Y_2	0	3/16	0	3/16
	1	1/16	3/16	1/16

Calcular su coeficiente de correlación e indicar si son independientes.

Solución: A partir de la tabla del enunciado se pueden obtener las distribuciones marginales de las dos variables sumando los valores por filas y por columnas:

		Y_1			Marginal Y_2
		-1	0	1	
	-1	1/16	3/16	1/16	5/16
Y_2	0	3/16	0	3/16	6/16
	1	1/16	3/16	1/16	5/16
Marginal Y_1		5/16	6/16	5/16	

Se observa de la tabla que las variables no cumplen que $P(Y_1 = a, Y_2 = b) = P(Y_1 = a)P(Y_2 = b)$, por tanto no son independientes. Sin embargo, el coeficiente de correlación lineal es 0. La media de las dos variables es cero:

$$E[Y_1] = E[Y_2] = \frac{5}{16} \times (-1) + \frac{6}{16} \times (0) + \frac{5}{16} \times (+1) = 0$$

y la covarianza

$$\begin{aligned}
\mathrm{Cov}(Y_1, Y_2) &= E[Y_1Y_2] \\
&= (-1)\times(-1)\frac{1}{16} + (-1)\times(0)\frac{3}{16} + (-1)\times(+1)\frac{1}{16} + \\
&\quad (0)\times(-1)\frac{3}{16} + (0)\times(0)\times 0 + (0)\times(+1)\frac{3}{16} + \\
&\quad (+1)\times(-1)\frac{1}{16} + (+1)\times(0)\frac{3}{16} + (+1)\times(+1)\frac{1}{16} = 0
\end{aligned}$$

Si la covarianza es cero, la correlación es cero. Sin embargo, Y_1 e Y_2 no son independientes aunque tienen coeficiente de correlación cero. Conclusión: dos variables con coeficiente de correlación igual a cero no tienen por qué ser independientes. Ahora bien, si dos variables son independeintes, entonces su coeficiente de correlación es cero.

• • •

23. La función de densidad conjunta de X e Y viene dada por

$$f(x, y) = xy, \qquad 0 < x < 1, \;\; 0 < y < 2$$

a) Obtener las funciones de densidad marginales y decir si X e Y son independientes.

b) Calcular $\mathrm{P}(X + Y < 1)$.

c) Calcular $\mathrm{P}(X + Y < 2)$.

Solución:

a)

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)dy = \int_0^2 xydy = x\frac{y^2}{2}\Big|_{y=0}^{y=2} = 2x, \quad 0 \le x \le 1.$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y)dx = \int_0^1 xydx = y\frac{x^2}{2}\Big|_{x=0}^{x=1} = \frac{1}{2}y, \quad 0 \le y \le 2.$$

Se observa que el producto de las dos marginales es igual a la distribución conjunta, por tanto son dos variables aleatorias independientes.

b) Ver figura 2.14 (izquierda)

$$\begin{aligned}
P(X + Y < 1) &= \int_0^1 \int_0^{1-y} xy\,dx\,dy = \int_0^1 y\left(\frac{x^2}{2}\Big|_0^{1-y}\right)dy \\
&= \frac{1}{2}\int_0^1 y(1-y)^2 dy = \frac{1}{24}
\end{aligned}$$

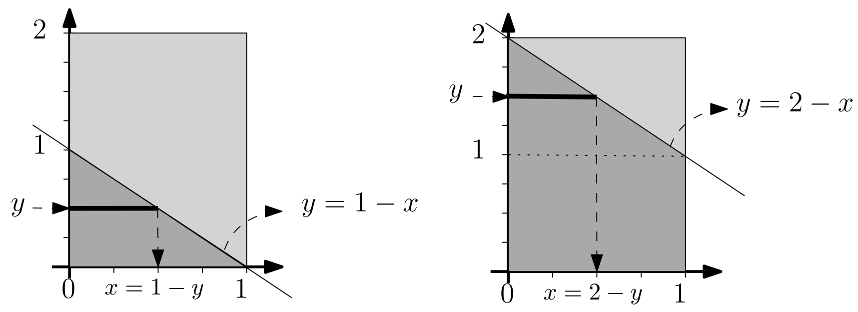


Figura 2.15: Recintos de integración (ejercicio 23)

.

c) Ver figura 2.14 (derecha)

$$\begin{aligned} P(X+Y<2) &= \int_0^1 \int_0^1 xy\,dx\,dy + \int_1^2 \int_0^{2-y} xy\,dx\,dy \\ &= \frac{1}{4} + \int_1^2 y\left(\frac{x^2}{2}\Big|_0^{2-y}\right) dy \\ &= \frac{1}{4} + \frac{1}{2}\int_1^2 y(2-y)^2 dy = \frac{17}{24} \end{aligned}$$

• • •

24. Un ordenador tarda un total de Y segundos en procesar un mensaje de correo electrónico, esta cantidad incluye el tiempo X durante el cual el mensaje está en la cola esperando a ser procesado ($Y \geq X$). La función de densidad conjunta de las variables aleatorias X, Y es

$$f_{X,Y}(x,y) = e^{-y}, \quad 0 \leq x \leq y < \infty$$

Calcular la probabilidad de que un mensaje haya estado menos de un segundo en la cola si el tiempo total que ha durado su procesamiento ha sido mayor que dos segundos.

Solución: La gráfica de la función de densidad conjunta se muestra en la figura 2.16. Se tiene que calcular

$$P(X \leq 1 | Y \geq 2) = \frac{P(X \leq 1, Y \geq 2)}{P(Y \geq 2)} \tag{2.3}$$

Para calcular $P(Y \geq 2)$ es necesario obtener previamente la función de densidad marginal de Y, (ver gráfica de la izquierda de la figura 2.17)

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)dx = \int_0^y e^{-y}dx = ye^{-y}, y \geq 0.$$

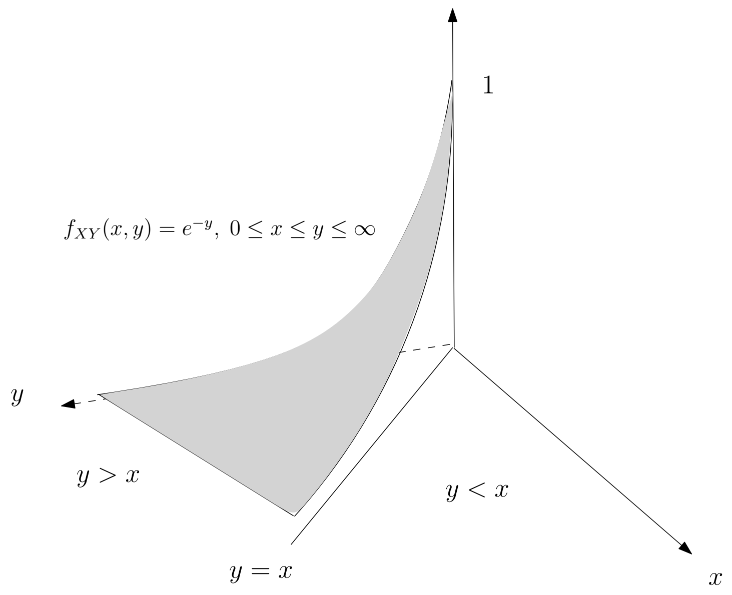


Figura 2.16: Función de densidad (ejercicio 24)
.

Por tanto,

$$P(Y \geq 2) = \int_2^{\infty} y e^{-y}\, dy$$

La integral se resuelve por partes, llamando $u = y$ y $dv = e^{-y}dy$, lo que implica que $du = dy$ y $v = -e^{-y}$ y sustituyendo

$$P(Y \geq 2) = -ye^{-y}\Big|_2^{\infty} + \int_2^{\infty} e^{-y}\, dy = 3e^{-2}$$

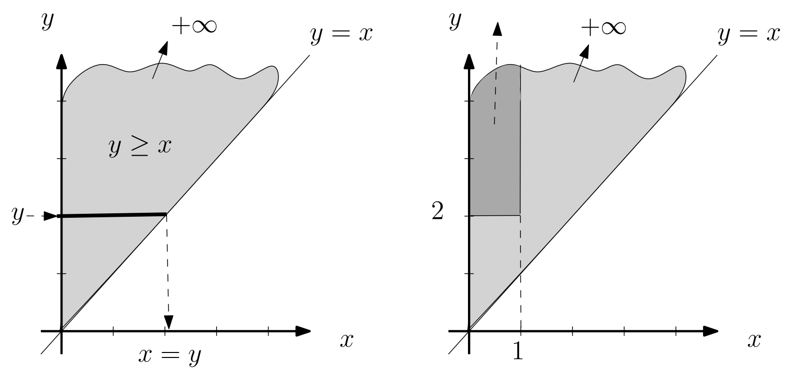


Figura 2.17: Recintos de integración (ejercicio 24)
.

• • •

Ahora vamos a calcular el numerador de la ecuación (2.3). $P(X \leq 1, Y \geq 2)$ se obtiene intengrando la función de densidad conjunta (ver gráfica de la derecha de la figura (2.17),

$$P(X \leq 1, Y \geq 2) = \int_0^1 \int_2^\infty e^{-y}\, dx\, dy = \left(\int_0^1 dx\right)\left(\int_2^\infty e^{-y} dy\right) = e^{-2}.$$

Ya hemos calculado el numerador y denominador de la ecuación (2.3), por tanto

$$P(X \leq 1 | Y \geq 2) = \frac{e^{-2}}{3e^{-2}} = \frac{1}{3}.$$

• • •

25. Sea X un valor elegido al azar de la distribución uniforme en el intervalo [0,1]. Se observa $X = x$ y a continuación se toma al azar otro valor Y de la distribución uniforme en el intervalo $[x, 1]$. Calcular la función de densidad marginal de Y.

    Solución: $f_X(x) = 1, \quad 0 \leq x \leq 1$. Se toma un valor al azar de X que denominaremos x. " A continuación se toma al azar otro valor Y de la distribución uniforme $[x, 1]$," significa que

$$f_{Y|X}(y|x) = \frac{1}{1-x}, \quad x \leq y \leq 1.$$

A partir de f_X y $f_{Y|X}$ podemos obtener la distribución conjunta de X e Y,

$$f_{XY}(x, y) = f_{Y|X}(y|x) f_X(x) = \underbrace{\frac{1}{1-x}}_{f_{Y|X}(y|x)} \times \underbrace{1}_{f_X(x)}, \quad 0 \leq x \leq y \leq 1$$

La función de densidad marginal de Y se obtienen como

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y) dx = \int_0^y \frac{1}{1-x} dx = -\log(1-y), \quad 0 \leq y \leq 1.$$

• • •

26. Sean X, Y, U y V variables aleatorias, demostrar que si $Y = U + V$, entonces

$$\mathrm{Cov}(X, Y) = \mathrm{Cov}(X, U) + \mathrm{Cov}(X, V).$$

Solución: Utilizando las propiedades del operador esperanza se tiene

$$\begin{aligned}
\mathrm{Cov}(X,Y) &= E[(X-\mu_X)(Y-\mu_Y)] \\
&= E[(X-\mu_X)((U+V)-(\mu_U+\mu_V))] \\
&= E[(X-\mu_X)((U-\mu_U)+(V-\mu_V))] \\
&= E[(X-\mu_X)(U-\mu_U)] + E[(X-\mu_X)(V-\mu_V)] \\
&= \mathrm{Cov}(X,U) + \mathrm{Cov}(X,V).
\end{aligned}$$

• • •

27. Sean dos variables aleatorias independientes X_1 y X_2 con la misma función de densidad $f(x) = e^{-x}, x \geq 0$

   a) Obtener la función de densidad conjunta.

   b) Obtener la función de densidad de $Y = X_1 + X_2$.

   c) Añadimos otra variable X_3 independiente e idénticamente distribuida que las anteriores, obtener la función de densidad de $Y_3 = X_1 + X_2 + X_3$.

   d) Dada n variables aleatorias independientes con la distribución anterior, se define $Y_n = X_1 + X_2 + \cdots + X_n$. Aplicando el procedimiento anterior de manera recurrente se obtiene que la función de densidad de Y_n es

$$f_{Y_n}(y) = \frac{y^{n-1}}{k(n)} e^{-y}, y \geq 0.$$

   Obtener $E[Y_n]$ y deducir la expresión de la función $k(n)$.

Solución:

a) $f_{X_1,X_2}(x_1,x_2) = f(x_1)f(x_2) = e^{-(x_1+x_2)}, \quad x_1 \geq 0, x_2 \geq 0$

b) Vamos a obtener la función de distribución de Y

$$\begin{aligned}
F_Y(y) = P(X_1+X_2 \leq y) &= \int_0^y \int_0^{y-x_2} e^{-(x_1+x_2)}\, dx_1\, dx_2 \\
&= \int_0^y e^{-x_2} \left(\int_0^{y-x_2} e^{-x_1} dx_1 \right) dx_2 = \int_0^y \left(e^{-x_2} - e^{-y} \right) dx_2 \\
&= -e^{-x_2} - x_2 e^{-y} \Big|_0^y = 1 - (1+y)e^{-y},\ y \geq 0.
\end{aligned}$$

Derivando la función de distribución obtenemos la función de densidad

$$f_Y(y) = \frac{d\,F_Y(y)}{d\,y} = ye^{-y}, y \geq 0.$$

c) Seguimos el procedimiento anterior, teniendo en cuenta que $Y_3 = X_1 + X_2 + X_3 = Y + X_3$, la función de densidad conjunta de X_3 e Y es

$$f_{X_3,Y}(x,y) = ye^{-(x+y)}, x \geq 0, y \geq 0,$$

la función de distribución de Y_3 es

$$\begin{aligned}
F_{Y_3}(t) = P(X_3 + Y \leq t) &= \int_0^t \int_0^{t-y} ye^{-(x+y)}\, dx\, dy \\
&= \int_0^t ye^{-y} \left(\int_0^{t-y} e^{-x} dx \right) dy = \int_0^t ye^{-y} \left(1 - e^{-(t-y)}\right) dy \\
&= \int_0^t \left(ye^{-y} - ye^{-t}\right) dy = 1 - e^{-t}(1 + t + \frac{t^2}{2}),\ t \geq 0
\end{aligned}$$

y derivando se obtiene la función de densidad

$$f_{Y_3}(t) = \frac{1}{2} t^2 e^{-t},\ t \geq 0$$

d) Como la integral entre 0 e ∞ de la función de densidad es igual a 1, tenemos que

$$k(n) = \int_0^\infty y^{n-1} e^{-y}\, dy \tag{2.4}$$

por otra parte sabemos que $E[Y_n] = E[X_1] + E[X_2] + \cdots + E[X_n] = n$, pues se puede comprobar fácilmente que $E[X_i] = 1$, por tanto

$$E[Y_n] = \int_0^\infty y \times \frac{y^{n-1}}{k(n)} e^{-y} dy = n,$$

y

$$\int_0^\infty y^n e^{-y}\, dy = nk(n)$$

que según (2.4) es igual a $k(n+1)$, es decir

$$k(n+1) = nk(n) \tag{2.5}$$

Sustituyendo $n = 1$ en la ecuación (2.4) se tiene $k(1) = 1$, y utilizando la recurrencia de la ecuación (2.5), tenemos que $k(2) = 1$, $k(3) = 2$, $k(4) = 6$ y

$$k(n) = (n-1)!$$

Capítulo 3

Modelos de probabilidad

1. Calcular la probabilidad de obtener al menos un seis al lanzar cuatro veces un dado.

   Solución: Llamando X al número de seises al lanzar cuatro veces un dado, se tiene que X es una variable aleatoria con distribución binomial de parámetros $n = 4$ y $p = 1/6$

$$\begin{aligned} P(\text{ Al menos un seis}) &= P(X > 0) = 1 - P(X = 0) \\ &= 1 - \binom{4}{0}\left(\frac{1}{6}\right)^0\left(\frac{5}{6}\right)^4 \\ &= 1 - \left(\frac{5}{6}\right)^4 \\ &= 0{,}518 \end{aligned}$$

• • •

2. Se lanza una pareja de dados 24 veces, calcular la probabilidad de por lo menos obtener una pareja de seises. (Este problema tiene interés histórico, fue resuelto por Pascal en el siglo XVIII para ayudar a un jugador llamado De Meré).

   Solución: El experimento consiste en repetir 24 veces el lanzamiento de dos dados. La probabilidad de obtener dos seises en un lanzamiento es

$$p = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$$

   Llamando Y al número de parejas de seises al lanzar $n = 24$ veces dos dados, se tiene que Y sigue una distribución Binomial con $n = 24$ y $p = 1/36$.

$$\begin{aligned} P(\text{ Al menos una pareja de seises}) &= P(Y > 0) = 1 - P(Y = 0) \\ &= 1 - \binom{24}{0}\left(\frac{1}{36}\right)^0\left(\frac{35}{36}\right)^{24} \\ &= 1 - \left(\frac{35}{36}\right)^{24} \\ &= 0{,}491 \end{aligned}$$

El caballero De Meré pensaba que el suceso sacar al menos un 6 al lanzar cuatro dados y el suceso sacar una pareja de 6 al lazar 24 veces dos dados eran sucesos equiprobables. No estaba muy desencaminado.

• • •

3. Una urna contiene 6 bolas, cuatro blancas y dos negras. Se extraen seis bolas con reposición, ¿Cuál es la probabilidad de que se obtengan 4 bolas blancas y dos negras? ¿Cuál es la probabilidad de que hayan extraído más bolas negras que blancas? ¿Cuál es la probabilidad de que hayan extraído más bolas blancas?

   Solución: Llamado X a la variable aleatoria número de bolas negras al sacar 6 bolas de la urna (con reposición), se tiene que X tiene distribución binomial con $n = 6$ y $p = 2/6 = 1/3$. Sacar dos bolas negras y cuatro blancas es el suceso $X = 2$,

$$P(X = 2) = \binom{6}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^4 = 0{,}3292$$

   El suceso " obtener más bolas negras que blancas" se puede poner como $X > 3$ y su probabilidad es

$$\begin{aligned}P(X > 3) &= P(X = 4) + P(X = 5) = P(X = 6)\\ &= \binom{6}{4}\left(\frac{1}{3}\right)^4\left(\frac{2}{3}\right)^2 + \binom{6}{5}\left(\frac{1}{3}\right)^5\left(\frac{2}{3}\right)^1 + \binom{6}{6}\left(\frac{1}{3}\right)^6\left(\frac{2}{3}\right)^0\\ &= 0{,}1001\end{aligned}$$

   Ahora el suceso " obtener más bolas blancas" será $X < 3$ y su probabilidad es

$$\begin{aligned}P(X < 3) &= P(X = 0) + P(X = 1) = P(X = 2)\\ &= \binom{6}{0}\left(\frac{1}{3}\right)^0\left(\frac{2}{3}\right)^6 + \binom{6}{1}\left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^5 + \binom{6}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^4\\ &= 0{,}6804\end{aligned}$$

• • •

4. De un lote que contiene 200.000 piezas se extraen al azar 250. Si el número de piezas defectuosas en la muestra es mayor que c se rechaza el lote y se devuelve al proveedor. Calcular c si se desea que la probabilidad de rechazar un lote con 3 % de piezas defectuosas sea 0.05. (Nota: suponer que se toma la muestra con remplazamiento para hacer los cálculos).

   Solución: El número de defectuosos en la muestra X sigue una distribución Binomial con parámetros $n = 250$ y $p = 0{,}03$. Se desea calcular c tal que $P(X > c) = 0{,}05$, es decir

$$P(X \leq c) = \sum_{k=0}^{c}\binom{250}{k}0{,}03^k 0{,}97^{250-k} = 0{,}95$$

El cálculo de c directamente a partir de la fórmula anterior no es fácil. Se puede ir tanteando. Más adelante veremos como resolverlo utilizando la distribución normal (ver ejercicio 41). Lo más sencillo es utilizar algún programa que incluya la distribución binomial. Por ejemplo, en `R`, el resultado $c = 12$ se obtiene utilizando la instrucción `qbinom(0.95,size=250,prob=0.03)`.

• • •

5. Si las llamadas telefónicas a una centralita siguen una distribución de Poisson de parámetro $\lambda = 3$ llamadas/cinco minutos, calcular la probabilidad de:

   a) Seis llamadas en cinco minutos.

   b) Tres llamadas en diez minutos.

   c) Más de 15 en un cuarto de hora.

   d) Dos en un minuto.

   Solución:

   a) Sea X la variable aleatoria número de llamadas en "cinco minutos", X tiene una distribución de Poisson de parámetro $\lambda = 3$, por tanto

   $$P(X = 6) = e^{-3}\frac{3^6}{6!} = 0{,}0504$$

   b) El número de llamadas en 10 minutos es una variable aleatoria Y con distribución de Poisson de parámetro $\lambda = 6$,

   $$P(Y = 3) = e^{-6}\frac{6^3}{3!} = 0{,}0892$$

   c) El número de llamadas en 15 minutos es una variable aleatoria U con distribución de Poisson de parámetro $\lambda = 9$,

   $$P(U > 15) = 1 - P(U \leq 15) = 1 - \sum_{u=0}^{15} e^{-9}\frac{9^u}{u!} = 0{,}0220$$

   d) El número de llamadas en un minuto es una variable aleatoria V con distribución de Poisson de parámetro $\lambda = 3/5$,

   $$P(V = 2) = e^{-3/5}\frac{(3/5)^2}{2!} = 0{,}0987$$

• • •

6. La variable aleatoria X tiene distribución exponencial con media 1. Obtener la función de distribución y la función de densidad de

$$W = aX^{1/b}, \quad a > 0, b > 0$$

Solución:

La función de distribución de W es por definición:

$$F_W(w) = P(W \leq w) = P(aX^{1/b} \leq w) = P(X \leq (\frac{w}{a})^b)$$

Si X tiene distribución exponencial de media 1, $f_X(x) = e^{-x}$ y la función de distribución es $F_X(x) = P(X \leq x) = 1 - e^{-x}$, $x \geq 0$, por tanto

$$F_W(w) = P(X \leq (\frac{w}{a})^b) = 1 - \exp{(-\frac{w}{a})^b}$$

La función de densidad se obtiene derivando:

$$f_W(w) = \frac{d}{dw}F_W(w) = \frac{b}{a}(\frac{w}{a})^{b-1}\exp{(-\frac{w}{a})^b}$$

• • •

7. El número de averías diarias de una máquina sigue una distribución de Poisson de media 0,4 averías. Calcular la probabilidad de que haya tres días sucesivos sin averías.

Solución:

El número de averías en 3 días es una variable aleatoria X con distribución de Poisson de parámetro $\lambda = 1{,}2$, por tanto

$$P(X = 0) = e^{-1{,}2} = 0{,}301$$

• • •

8. A un puesto de servicio llegan de manera independiente, por término medio, 10 clientes/hora. Calcular la probabilidad de que lleguen 8 clientes en la próxima media hora sabiendo que en la última hora llegaron 14 clientes, y que la variable aleatoria *número de clientes que llegan en un hora* siguen una distribución de Poisson.

Solución:

En un proceso de Poisson los sucesos que corresponden a intervalos que no se solapan son independientes. La variable aleatoria X *número de clientes en la próxima media hora* es independiente de la variable aleatoria Y *número de clientes en la hora anterior*. Como X es una variable de Poisson de parámetro $\lambda = 5$, se tiene que:

$$P(X = 8) = \exp{(-5)}\frac{5^8}{8!} = 0{,}0653$$

• • •

9. Sean X e Y dos variables independientes con distribución geométrica de parámetros 0.4 y 0.6 respectivamente. Calcular $P(X+Y=3)$.

Solución:

$$\begin{aligned} P(X+Y=3) &= P(X=1, Y=2) + P(X=2, Y=1) \\ &= P(X=1)P(Y=2) + P(X=2)P(Y=1) \\ &= 0{,}4 \times (0{,}4 \times 0{,}6) + 0{,}6 \times (0{,}6 \times 0{,}4) = 0{,}24 \end{aligned}$$

• • •

10. En una planta industrial (ver 3.1) dos bombas B_1 y B_2 en paralelo conducen agua desde un pozo a una depuradora D, y posteriormente otras dos bombas B_3 y B_4, también en paralelo, la trasladan a un depósito como indica la figura.

Los tiempos de vida de la depuradora y de las bombas son variables aleatorias independientes con distribución exponencial, siendo 20 mil horas la vida media de la depuradora y 30 mil horas la de cada bomba.

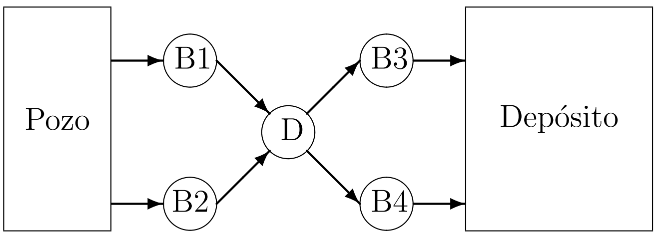


Figura 3.1: Planta industrial de depuración

a) Calcular la probabilidad de que llegue agua al depósito después de 20 mil horas de funcionamiento.

b) Calcular la probabilidad de que una depuradora que ha trabajado T horas falle antes de las mil horas siguientes. ¿Es razonable que para evitar fallos de la depuradora se renueve ésta cada 20 mil horas? ¿Por qué?

Solución:

a) Sea X el tiempo de funcionamiento de una bomba e Y el tiempo de funcionamiento de la depuradora. Las funciones de densidad de X e Y son respectivamente:

$$f_X(x) = \frac{1}{30}e^{-x/30}, \qquad f_Y(y) = \frac{1}{20}e^{-y/20}$$

La función de distribución exponencial aplicada a la bomba es $F_X(x) = P(X \leq x) = 1 - e^{-x/30}$ y por tanto $P(X \geq x) = e^{-x/30}$. La probabilidad de que una bomba funcione más de 20 mil horas es

$$P(X \geq 20) = e^{-20/30} = 0{,}5134$$

y la probabilidad de que la depuradora dure más de 20 mil horas es

$$P(Y \geq 20) = e^{-20/20} = 0{,}3679$$

Para analizar la fiabilidad del sistema vamos a considerar que está formado por tres grupos en serie: el grupo $G1$ de bombas, la depuradora D y el grupo $G2$ de bombeo. Para que el agua vaya del pozo al depósito, deben funcionar los tres grupos (ver figura 3.2).

El grupo $G1$ funciona después de 20 mil horas si funciona al menos una bomba de las dos $B1$ o $B2$. La probabilidad de que fallen las dos es $(1 - 0{,}5134)^2 = 0{,}2368$ y que funcione al menos una es $1 - (1 - 0{,}5134)^2 = 0{,}7632$ El grupo $G2$ es idéntico, tiene la misma probabilidad de funcionar después de 20 mil horas. Por tanto la probabilidad de que el conjunto funcione después de 20 mil horas es

$$0{,}7632 \times 0{,}3679 \times 0{,}7632 = 0{,}2143$$

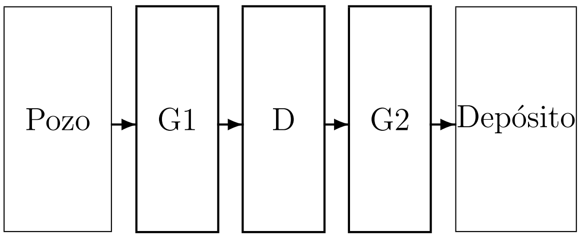


Figura 3.2: Esquema de la planta industrial

b) Vamos a calcular la probabilidad de que dure más de $T + 1$ horas sabiendo que ha durado más T, es decir

$$\begin{aligned} P(X > T + 1 | X > T) &= \frac{P(X > T, X > T + 1)}{P(X > T)} \\ &= \frac{P(X > T + 1)}{P(X > T)} = \frac{e^{-\lambda(T+1)}}{e^{-\lambda T}} \\ &= e^{-\lambda} = P(X > 1) \end{aligned}$$

La probabilidad de que dure mil horas más, después de llevar funcionando T horas, es la misma que la probabilidad de que una depuradora nueva dure más de mil horas. Es decir que -según ésto- el dispositivo no sufre "desgaste". Es habitual referirse a esta propiedad de la distribución exponencial con la afirmación: ”la distribución exponencial no tiene memoria”. La probabilidad pedida es

$$\begin{aligned} P(X < T+1|X > T) &= 1 - P(X > T+1|X > T) \\ &= 1 - P(X > 1) = 1 - e^{-1/20} = 0{,}0488 \end{aligned}$$

Si el comportamiento de la depuradora es el anterior, no tiene sentido renovar una depuradora antes de que falle, porque en cada instante funciona como si estuviera nueva (su probabilidad de fallo no cambia).

• • •

11. Un laboratorio de análisis realiza pruebas de sangre para detectar la presencia de un tipo de virus. Se sabe que una de cada 100 personas es portadora del virus. Se va a realizar un estudio en un colegio, para abaratar las pruebas se realiza un análisis combinado que consiste en: En lugar de analizar la sangre de cada individuo, se toman las muestras de 50 y se analiza la mezcla. Si el resultado del análisis es negativo, se concluye que los 50 individuos están sanos. Si el análisis es positivo, se repite a cada persona de manera individual. El análisis es infalible.

   a) Determinar el número esperado de pruebas (análisis) que se tendrá que realizar si se sigue este tipo de estrategia.

   b) ¿Cuál es la probabilidad de que un individuo determinado sea portador del virus, si el resultado del análisis realizado a su grupo de 50 ha resultado positivo?

   Solución:

   a) Sea X el número de individuos con el virus en una muestra de 50. X tiene distribución binomial con parámetros $n = 50$ y $p = 1/100$. La probabilidad de que los 50 individuos estén sanos es $0{,}99^{50} = 0{,}605$ y la probabilidad de que uno o más tengan el virus es igual a 0.395. Si se sigue este procedimiento, el número esperado de pruebas que hay que hacer (por cada 50) es

   $$1 \times 0{,}605 + 51 \times 0{,}395 = 20{,}7$$

   b) Sea A el suceso un individuo del grupo tiene el virus y G el grupo ha dado positivo. Según el enunciado $P(A) = 0{,}01$ y hemos calculado $P(G) = 0{,}395$, aplicando el teorema de Bayes:

   $$\begin{aligned} P(A|G) &= \frac{P(A \cap G)}{P(G)} \\ &= \frac{P(G|A)P(A)}{P(G)} \\ &= \frac{1 \times (1/100)}{1 - 0{,}99^{50}} \\ &= 0{,}0253 \end{aligned}$$

Se ha utilizado que $P(G|A) = 1$, y significa que si A que pertenece al grupo tiene el virus, el grupo G tiene el virus con probabilidad 1. Si el grupo tiene el virus, la probabilidad de que la tenga un individuo en concreto es al menos $1/50 = 0.02$, la probabilidad calculada es ligeramente superior 0.0253, y tiene en cuenta que puede haber más de un individuo en el grupo con el virus.

• • •

12. De un lote con una proporción de piezas defectuosas p, se extraen piezas con reposición hasta que se observa la $k-$ésima defectuosa. Obtener la distribución de probabilidad de la variable aleatoria X *número total de piezas observadas.*

Solución:

Por ejemplo, el suceso "tener que sacar 8 piezas para observar 3 defectuosas ($k = 3$)", implica que en las primeras 7 extracciones hay 2 defectuosas y que la octava es defectuosa. La última pieza es siempre defectuosa en este problema.

De forma general, la variable tomar el valor x si en las $x - 1$ primeras piezas hay $k - 1$ defectuosas y la última es defectuosa. La probabilidad de observar $k - 1$ en $x - 1$ extracciones se calcula con la distribución binomial. Por consiguiente $P(X = x) = P($ sacar x-1 defectuosas en k-1$) \times P($la última defectuosa$)$,

$$P(X = x) = \binom{x-1}{k-1} p^{k-1}(1-p)^{x-k} \times p$$

$$P(X = x) = \binom{x-1}{k-1} p^{k}(1-p)^{x-k}, \quad x = k, k+1, k+2, ...$$

A esta distribución se denomina *binomial negativa.*

• • •

13. La función de densidad de una variable aleatoria X viene dada por la expresión

$$f(x) = \begin{cases} x/8, & \text{si } 0 \leq x \leq 4 \\ 0, & \text{en el resto} \end{cases}$$

Se generan secuencialmente valores de esta variable. ¿Cuántos valores de X habrá que generar por término medio hasta obtener un valor mayor que 3?

Solución:

La probabilidad de obtener un valor mayor que 3 es

$$P(X > 3) = 1 - \int_0^3 \frac{x}{8} dx = \frac{9}{16}$$

Sea Y la variable aleatoria número de valores generados hasta que salga un resultado mayor que 3, se aprecia que Y tiene distribución geométrica de parámetro $p = 9/16$. La media de Y es $E[Y] = 1/p = 16/9 = 1{,}78$

• • •

14. Una pareja decide tener hijos hasta el nacimiento de la primera niña. Calcular la probabilidad de que tengan más de 4 hijos. (Supóngase $P(niño) = P(niña) = 0{,}5$)

Solución:

X el número de hijos es una variable aleatoria con distribución geométrica $P(X = x) = (1/2)^x, x = 1, 2, 3, ...,$ y

$$P(X > 4) = \sum_{5}^{\infty} \left(\frac{1}{2}\right)^x = \frac{(1/2)^5}{1-(1/2)} = \left(\frac{1}{2}\right)^4 = \frac{1}{16}.$$

• • •

15. La distancia D entre dos vehículos consecutivos en una autopista sigue una distribución exponencial con media 200 metros. ¿Cuál es la probabilidad de que en un tramo de 1 km haya exactamente 5 vehículos?

Solución:

La variable aleatoria D tiene distribución exponencial

$$f_D(x) = \lambda e^{-\lambda x}, \quad x > 0$$

la media de la variable aleatoria es $E[D] = 1/\lambda = 200m$, por tanto $\lambda = 1/200$ vehículos por metro. Sea Y la variable aleatoria *número de vehículos en 1 kilómetro*, tiene distribución de Poisson de media $\lambda_1 = 1000 \times \dfrac{1}{200} = 5$ vehículos por kilómetro, por tanto:

$$P(Y = 5) = e^{-5}\frac{5^5}{5!} = 0{,}175$$

• • •

16. Ricardo es un pescador experto que ha comprobado, después de una larga experiencia practicando su deporte favorito, que el número de peces capturados por la mañana puede ser representado por una variable aleatoria de Poisson de media 3 peces a la hora. Quiere ir a pescar el sábado próximo, si empieza a las 7 de la mañana, ¿cuál es la probabilidad de que capture el primer pez antes de las 7 h. 15 min.? ¿Cuál es la probabilidad de que capture 5 peces durante dos horas de pesca?

Solución:

Sea X el número de peces capturados en 15 minutos. X tiene distribución de Poisson de media 3/4. Capturar el primer pez antes de las 7:15 significa que $X > 0$ por tanto

$$P(X > 0) = 1 - e^{-3/4} = 0{,}5276$$

El número de peces en 2 horas es una variable aleatoria con distribución de Poisson de media 6, y la probabilidad de captura 5 peces es:

$$P(Y = 5) = e^{-6}\frac{6^5}{5!} = 0{,}1606$$

• • •

17. La variable aleatoria T representa la duración de vida de un componente electrónico. En teoría de la fiabilidad la probabilidad de que un componente falle en el instante t sabiendo que ha durado hasta t se denomina *tasa de fallo* y se representa por $\lambda(t)$, siendo su valor en función de t

$$\lambda(t) = \frac{f(t)}{1 - F(t)},$$

donde f y F son, respectivamente, las funciones de densidad y de distribución de la variable aleatoria T. Obtener la tasa de fallo en caso que T sea una variable aleatoria exponencial de media 1000 horas e interpretar el resultado.

Solución:

La función de densidad de T es

$$f_T(t) = \frac{1}{1000} e^{-t/1000},\ t > 0$$

y la función de distribución

$$F_T(t) = 1 - e^{-t/1000}.$$

Sustituyendo en la definición de $\lambda(t)$

$$\lambda(t) = \frac{f(t)}{1 - F(t)} = \frac{\frac{1}{1000} e^{-t/1000}}{e^{-t/1000}} = \frac{1}{1000}$$

que se interpreta que la tasa de fallo es constante e igual a un fallo cada mil horas. Lo habitual es que la tasa de fallo sea una función creciente con t, cuanto más antiguo es el componente mayor es la probabilidad de que falle. La distribución exponencial tiene una tasa de fallo constante, que indica que el dispositivo no envejece. Esta propiedad tiene que ver con la falta de memoria de la distribución exponencial.

• • •

18. Un examen consiste en 25 cuestiones. En cada cuestión, el alumno debe elegir entre 5 soluciones propuestas, de las que una (y sólo una) es cierta. El número mínimo de respuestas correctas que debe tener un alumno para aprobar es a. El profesor decide fijar a con el siguiente criterio: que la probabilidad de aprobar para un alumno que conteste todas las cuestiones al azar sea menor de 0.05. Obtener a. (Una cuestión es respondida al azar si cada uno de los cinco resultados propuestos tiene la misma probabilidad de ser escogido).

Solución:

El número de aciertos X de una persona que conteste al azar es una variable aleatoria con distribución binomial de parámetros $n = 25$ y $p = 1/5$. Se desea obtener a tal que

$$P(X > a) = 0{,}05$$

Se trata de obtener el mínimo valor a que cumple

$$\sum_{x=0}^{a} \binom{25}{x} p^x (1-p)^{25-x} = 0{,}95$$

La solución es $a = 8$ (en R se obtiene utilizando la instrucción qbinom(.95,n=25,p=0.2)). Utilizando la aproximación normal, X tiene media $25 \times 0{,}2 = 5$ y desviación típica $\sqrt{25 \times 0{,}2 \times 0{,}8} = 2$

$$P(X \leq a) = P(\frac{X-5}{2} \leq \frac{a-5}{2}) = 0{,}95$$

En las tabla de la N(0,1), $P(Z \leq 1{,}65) = 0{,}95$, por tanto

$$\frac{a-5}{2} = 1{,}65$$

y a

$$a = 5 + 1{,}65 \times 2 = 8{,}3$$

Se puede comprobar con el cálculo exacto si conviene utilizar $a = 8$ o $a = 9$, en este caso la solución es $a = 8$.

• • •

19. Obtener la función de densidad de una variable aleatoria χ^2 con un grado de libertad. (Si $X \rightsquigarrow N(0,1), Y = X^2$ es una χ^2_1.)

Solución:

$$F_Y(y) = P(Y \leq y) = P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y})$$

Llamando

$$\Phi(u) = \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx,$$

se tiene que

$$F_Y(y) = \Phi(\sqrt{y}) - \Phi(-\sqrt{y})$$

La función de densidad se obtiene derivando:

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{1}{2\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-y/2} + \frac{1}{2\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-y/2}$$

$$f_Y(y) = \frac{1}{\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-y/2} \quad y > 0$$

• • •

20. Dada una variable aleatoria X, cuya distribución es $N(0, \sigma^2)$, calcular la mediana de la variable $Y = |X|$.

Solución:

$$F_Y(y) = P(Y \leq y) = P(|X| \leq y) = P(-y \leq X \leq y)$$

La mediana es el valor m que cumple

$$F_Y(m) = 0{,}5,$$

por tanto se trata de obtener el valor m tal que (ver figura 3.3)

$$P(-m \leq X \leq m) = 0{,}5$$

o lo que es lo mismo

$$P(X \leq m) = 0{,}75$$

Estandarizando

$$P(X \leq m) = P(\frac{X}{\sigma} \leq \frac{m}{\sigma}),$$

y teniendo en cuenta que si Z es una N(0,1), $P(Z \leq 0{,}6745) = 0{,}75$ se tiene que

$$\frac{m}{\sigma} = 0{,}6745, \quad \Rightarrow \quad m = 0{,}6745\,\sigma.$$

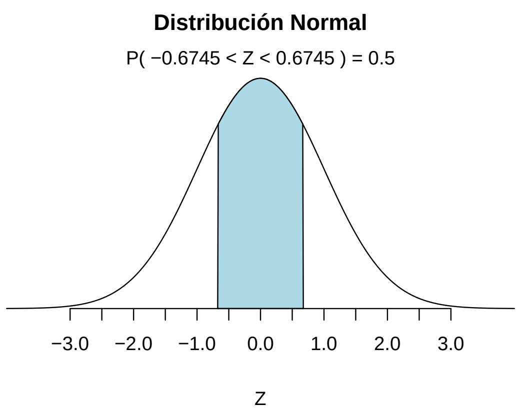


Figura 3.3: Normal estándar con intervalo central con 50 % probabilidad

• • •

21. La longitud L en milímetros de las piezas fabricadas en un proceso es una variable aleatoria que se distribuye según una $N(32, 0{,}3)$, considerándose aceptables aquellas cuya medida se encuentra dentro del intervalo (31.1, 32.6).

    a) Calcular la probabilidad de que una pieza elegida al azar sea aceptable.

    b) Si se toma al azar una muestra de tres piezas, ¿cuál es la probabilidad de que la primera y la tercera sean aceptables y la segunda no lo sea?

    c) ¿Cuál es la probabilidad de que en una muestra de tamaño 3 al menos una sea aceptable?

    d) Las piezas se embalan en lotes de 500. Calcular la probabilidad de que un lote tenga más de 15 defectuosas.

Solución:

a) La pieza es aceptable si $31{,}1 \leq L \leq 32{,}6$,

$$\begin{aligned} P(31{,}1 \leq L \leq 32{,}6) &= P\left(\frac{31{,}1-32}{0{,}3} \leq \frac{L-32}{0{,}3} \leq \frac{32{,}6-32}{0{,}3}\right) \\ &= P(-3 \leq Z \leq +2) \\ &= P(Z \leq +2) - P(Z \leq -3) \\ &= 0{,}9772 - 0{,}00135 = 0{,}976 \end{aligned}$$

Ver figura (3.4)

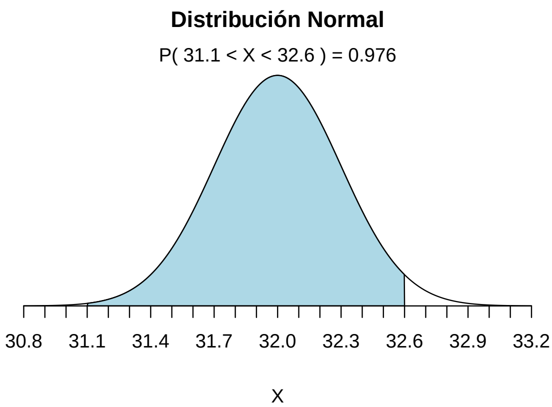


Figura 3.4: Calculo de probabilidad del ejercicio 21

b) Podemos escribir el suceso *"la primera y la tercera sean aceptables y la segunda no"* como $A_1 \cap D_2 \cap A_3$, por ser independientes:

$$P(A_1 \cap D_2 \cap A_3) = P(A_1)P(D_2)P(A_3) = 0{,}976 \times (1-0{,}976) \times 0{,}976 = 0{,}023$$

c) El número de aceptables en una muestra de tamaño 3 es una variable aleatoria X con distribución binomial de parámetros $n = 3$ y $p = 0{,}976$,

$$P(X \geq 1) = 1 - P(X = 0) = 1 - 0{,}0241^3 = 0{,}999986$$

d) El número de defectuosas Y en el lote de tamaño 500 es una variable aleatoria binomial con parámetros n=500 y p= 0.0241,

$$P(Y \leq 15) = \sum_{y=0}^{15} \binom{500}{y} p^y (1-p)^{500-y} = 0{,}8434$$

Utilizando la aproximación normal, $Y \sim N(np, \sqrt{np(1-p)})$, con $np = 12{,}05$ y $\sqrt{np(1-p)} = 3{,}43$, se tiene que utilizando la corrección por continuidad:

$$P(Y \leq 15{,}5) = P\left(\frac{Y - 12{,}05}{3{,}43} \leq \frac{15{,}5 - 12{,}05}{3{,}43}\right) = P(Z < 1) = 0{,}8413$$

La probabilidad de que haya más de 15 defectuosas es $P(Y \geq 16) = 1 - 0{,}8413 = 0{,}1587$

• • •

22. Un concesionario de automóviles recibe pedidos de un modelo según un proceso de Poisson de media 2 vehículos por semana. Los pedidos al fabricante se deben realizar con una antelación mínima de un mes, de forma que el concesionario pide en cada mes los vehículos que necesita para el mes siguiente. ¿Cuántos automóviles disponibles ha de tener a principios de un mes para satisfacer con probabilidad igual o mayor que 0.95 la demanda mensual? (Se considera que el mes tiene cuatro semanas).

Solución:

El número de vehículos Y demandados al mes es una variable aleatoria de media $\lambda = 8$. Se trata de obtener el valor a tal que $P(Y \leq a) = 0{,}95$,

$$P(Y \leq a) = e^{-8} \sum_{y=0}^{a} \frac{8^y}{y!} = 0{,}95$$

La solución $a = 13$ se obtiene añadiendo sumandos hasta que se alcanza el valor 0.95. En `R` se obtiene con la instrucción `a = qpois(0.95,8)`.

Otra forma de hacerlo es utilizando la aproximación normal, $Y \sim N(8, \sqrt{8})$ y

$$P(Y \leq a) = P\left(\frac{Y-8}{\sqrt{8}} \leq \frac{a-8}{\sqrt{8}}\right) = P\left(Z \leq \frac{a-8}{\sqrt{8}}\right) = 0{,}95,$$

siendo Z la normal estándar. Teniendo en cuenta que $P(Z \leq 1{,}65) = 0{,}95$, se tiene que

$$\frac{a-8}{\sqrt{8}} = 1{,}65 \quad \Rightarrow a = 8 + 1{,}65 \times \sqrt{8} = 12{,}7 \simeq 13$$

• • •

23. Si la probabilidad de que un disparo impacte una diana es 0.0001, ¿cuál es la probabilidad de impactar en la diana 4 o más veces en 50000 disparos? Da un resultado numérico empleando la aproximación que consideres más adecuada. Se supone independencia.

    Solución:

    X la variable aleatoria número de impactos en el blanco tras 50000 disparos sigue una distribución binomial de parámetros $n = 50000$ y $p = 0{,}0001$. Cuando p es pequeño y n es grande, X se puede aproximar por la distribución de Poisson de media $\lambda = 50000 \times 0{,}0001 = 5$, entonces

    $$P(X \leq 3) = e^{-5}\left(\frac{5^0}{0!} + \frac{5^1}{1!} + \frac{5^2}{2!} + \frac{5^3}{3!}\right) = 0{,}265$$

    La probabilidad pedida es $P(X \geq 4) = 1 - 0{,}265 = 0{,}735$.

    También se puede hacer utilizando la aproximación por la normal.

• • •

24. En una urna hay 20 bolas, 5 son negras y 15 blancas. Se extraen 4 sin reposición y se define la variable aleatoria Y como el número de bolas negras.

    a) Demostrar que la distribución de probabilidad de Y es

    $$P(Y = k) = \frac{\binom{5}{k}\binom{15}{4-k}}{\binom{20}{4}}, \quad k = 0, 1, 2, 3, 4$$

    b) Generalizar el resultado anterior si en la urna hay M bolas blancas, N bolas negras, se extraen K bolas e interesa conocer la distribución de probabilidad de Y número de bolas negras. La extracción es sin reposición. (Esta distribución se denomina *Hipergeométrica* y se utiliza en los problemas de muestreo sin reposición)

    c) Plantear el problema del número de aciertos en la lotería primitiva como un caso particular de la distribución anterior. Aplicarlo para el caso en el que el participante tacha (elige) 6 números y en el caso en el que tacha 8 números.

    Solución:

    a) Imagina que las bolas están numeradas del 1 al 20. Las 15 primeras son blancas y las restantes 5 negras. Con un conjunto de 20 bolas se pueden formar $\binom{20}{4} = 4845$ subconjuntos diferentes de cuatro bolas. En este cálculo el orden no importa y no se puede repetir número, porque la extracción es sin reposición.

En la tabla (3.1) se muestran las distintas composiciones en función del número de bolas blancas y negras. Hay 1365 combinaciones que tienen todas las bolas blancas, es decir con las 15 bolas blancas se pueden formar un total de $\binom{15}{4} =$ 1365 combinaciones distintas. Otro ejemplo, el número de combinaciones que tienen dos bolas blancas y dos bolas negras es $\binom{15}{2}\binom{5}{2} = 1050$, se obtiene multiplicando $\binom{15}{2}$ que es el número de formas de extraer 2 bolas entre las 15 blancas de la urna por $\binom{5}{2}$ el número de formas de extraer 2 bolas negras entre las 5 negras de las urna. La suma de las 5 filas de la tabla (3.1) es igual al número total de combinaciones 4845. Si se extraen 4 bolas al azar, cualquiera de estas 4845 combinaciones tiene la misma probabilidad de salir. Llamando Y a la variable aleatoria *número de bolas negras*, de la observación de la tabla se concluye que

$$P(Y = k) = \frac{\binom{5}{k}\binom{15}{4-k}}{\binom{20}{4}}, \quad k = 0, 1, 2, 3, 4$$

Tabla 3.1: Opciones con 4 bolas

Negras	Blancas	Número
0	4	$\binom{15}{4}\binom{5}{0} = 1365$
1	3	$\binom{15}{3}\binom{5}{1} = 2275$
2	2	$\binom{15}{2}\binom{5}{2} = 1050$
3	1	$\binom{15}{1}\binom{5}{3} = 150$
4	0	$\binom{15}{0}\binom{5}{4} = 5$
Total		$\binom{20}{4} = 4845$

b) El resultado se puede generalizar al caso de una urna con M bolas blancas y N bolas negras de donde se extráen K bolas sin reposición. X el número de bolas negras tiene la siguiente distribución de probabilidad:

$$P(X = x) = \frac{\binom{N}{x}\binom{M}{K-x}}{\binom{M+N}{K}}, \quad x = max(0, K - M), ..., min(K, N)$$

Si el número de bolas que se sacan K es menor que el número de bolas negras $K \leq N$ y también menor del número de bolas blancas $K \leq M$, los valores que puede tomar x van de 0 a K. Si el número de bolas que se sacan es mayor que el número de bolas blancas o que el número de bolas negras, los posibles valores de x deben tener en cuenta esas restricciones.

c) Esta distribución tiene muchas aplicaciones a juegos de azar. Por ejemplo la probabilidad de obtener un poker de ases al coger cinco cartas de una baraja francesa se obtiene utilizando la fórmula anterior. Otro caso es la lotería primitiva. Un participante elige 6 números entre 49. Podemos considerar que los 6 números elegidos son bolas negras y los restantes 43 son bolas blancas, la probabilidad de X el número de aciertos es:

$$P(X = x) = \frac{\binom{6}{x}\binom{43}{6-x}}{\binom{49}{6}}, \quad x = 0, 1, 2, ..., 6$$

En este caso $N = 6$, $M = 43$ y $K = 6$. Si el participante elige 8 números, los parámetros ahora serían $N = 8$, $M = 41$ y $K = 6$

$$P(X = x) = \frac{\binom{8}{x}\binom{41}{6-x}}{\binom{49}{6}}, \quad x = 0, 1, 2, ..., 6$$

En la tabla 3.2 se proporcionan los valores numéricos para estas distribuciones para diferentes casos (número de bolas negras o números elegidos).

Tabla 3.2: Probabilidad de aciertos en la Primitiva en función del núm. elegidos.

Aciertos	6	7	8	9	10
0	0.43596498	0.37513265	0.32154227	0.27448731	0.23331421
1	0.41301945	0.42582626	0.42872303	0.42349470	0.41173096
2	0.13237803	0.16808931	0.20277441	0.23527483	0.26468419
3	0.01765040	0.02873322	0.04268935	0.05934861	0.07842494
4	0.00096862	0.00215499	0.00410475	0.00702812	0.01112786
5	0.00001845	0.00006307	0.00016419	0.00036042	0.00070281
6	0.00000007	0.00000050	0.00000200	0.00000601	0.00001501
Apuestas	1	7	28	84	210

• • •

25. Para controlar la calidad de un proceso textil se cuenta el número de defectos que aparecen en la tela fabricada. Según el fabricante, cuando el proceso funciona correctamente el número de defectos en una bobina de 100 metros cuadrados es una variable aleatoria de Poisson con media 4. Se ha instalado un equipo de visión artificial para realizar el recuento que permite inspeccionar 900 m^2 de tela cada hora. ¿Cuál es la probabilidad de que aparezcan más de 50 defectos en una hora si el proceso funciona bien?

Solución:

Si el número de defectos cada 100 m^2 tiene distribución de Poisson de media 4, el número de defectos cada 900 m^2 tiene distribución de Poisson de media $\lambda = 9 \times 4 = 36$. Esta variable Y se puede aproximar por la distribución normal, $Y \sim N(\lambda, \sqrt{\lambda})$.

$$P(Y \geq 50) = P\left(\frac{Y-36}{\sqrt{36}} \geq \frac{50-36}{\sqrt{36}}\right) = P\left(Z \geq 2{,}33\right) = 0{,}0099$$

• • •

26. Un compañía compra *chips* para montar en placas de ordenadores clónicos. Una empresa de reciclado le ofrece lotes de 10.000 chips a precios muy ventajosos pero con un porcentaje de defectuosos alto, alrededor del 10 %. Para realizar el control de calidad de los lotes recibidos está considerando dos alternativas: (a) Tomar 100 unidades al azar y rechazar el lote si existen más de 15 defectuosas. (b) Tomar 100 unidades al azar, dividirlas en 10 grupos y si en algún grupo hay más de una pieza defectuosa rechazar el lote.¿Cuál es la probabilidad de rechazar un lote con el 10 % de chips defectuosos para cada uno de los métodos? ¿Cuál es la probabilidad de aceptar un lote con el 20 % de chips defectuoso para cada método?

Solución:

a) X la variable aleatoria *número de defectuosas* tiene distribución binomial. Si el lote tiene 10 % de piezas defectuosas ($p = 0{,}1$), la probabilidad de aceptarlo es

$$P(X \leq 15) = \sum_{x=0}^{15} \binom{100}{x} 0{,}1^x (1-0{,}1)^{100-x} = 0{,}9601$$

La probabilidad de rechazarlo es $P(X \geq 16) = 0{,}0399$. Se han realizado los cálculos con `R`, con la instrucción `pbinom(15,100,0.1`.

La manera de hacerlo si no se tiene la posibilidad de un programa infromático es utilizando la aproximación normal: $X \sim N(np, \sqrt{np(1-p)})$, con $np = 10$ y $\sqrt{np(1-p)} = 3$, se tiene que utilizando la corrección por continuidad:

$$P(X \leq 15{,}5) = P\left(\frac{X-10}{3} \leq \frac{15{,}5-10}{3}\right) = P(Z < 1{,}83) = 0{,}9664$$

Con esta aproximación, la probabilidad de rechazar el lotes es 0.0336.

En el segundo método, dividir la muestra en grupos de 10 piezas, la variable Y número de piezas defectuosa en el grupo de 10 tiene distribución binomial, la probabilidad de que en una muestra haya 0 o 1 piezas defectuosas es:

$$P(Y \leq 1) = P(Y=0) + P(Y=1) = \binom{10}{0} 0{,}1^0 (1-0{,}1)^{10} + \binom{10}{1} 0{,}1^1 (1-0{,}1)^9 = 0{,}7361$$

Para aceptar el lote, los 10 grupos deben tener menos de 2 piezas defectuosas, por tanto la probabilidad de aceptar el lote es $P(Y \leq 1)^{10} = 0{,}7361^{10} = 0{,}0467$ y la probabilidad de rechazarlo 0.9533.

Los dos métodos son radicalmente distintos. La probabilidad de rechazar un lote con el 10 % de piezas defectuosas del primero es 0.0336 y la del segundo es 0.9533.

b) La probabilidad de aceptar un lote con 20 % de piezas defectuosas por el método (a) es utilizando la distribución binomial

$$P(X \leq 15) = \sum_{x=0}^{15} \binom{100}{x} 0{,}2^x (1 - 0{,}2)^{100-x} = 0{,}1285$$

y utilizando la aproximacií normal

$$P(X \leq 15{,}5) = P\left(\frac{X - 20}{4} \leq \frac{15{,}5 - 20}{4}\right) = P(Z < -1{,}125) = 0{,}1303$$

Para el segundo método (b),

$$P(Y \leq 1) = \binom{10}{0} 0{,}2^0 (0{,}8)^{10} + \binom{10}{1} 0{,}2^1 (0{,}8)^9 = 0{,}3758$$

por tanto la probabilidad de aceptar el lote es $P(Y \leq 1)^{10} = 0{,}3758^{10} = 5{,}62e - 5$

También los dos métodos son diferentes en este caso.

• • •

27. Una compañía para comprobar la calidad de ciertos lotes de 30000 piezas realiza el siguiente control: toma una muestra al azar de 300 piezas y si tiene 15 o más piezas defectuosas rechaza el lote, aceptándolo en caso contrario. La compañía cada mes aplica este control a 200 lotes, ¿cuál es el número esperado de lotes rechazados si todos los lotes de un mes tienen exáctamente un 4 % de piezas defectuosas?

Solución:

Sea X el número de piezas defectuosas en una muestra de tamaño $n = 300$. Si el lote tiene una proporción $p = 0{,}04$ de piezas defectuosas, entonces X tiene distribución binomial. Se acepta el lote si $X \leq 14$ y la probabilidad es

$$P(X \leq 14) = \sum_{x=0}^{14} \binom{300}{x} 0{,}04^x \, 0{,}96^{300-x} = 0{,}7757$$

Este cálculo se ha realizado con `R` mediante la instrucción `pbinom(14,300,0.04)`.

Utilizando la aproximación normal, $X \sim N(300 \times 0{,}04, \sqrt{300 \times 0{,}04 \times 0{,}96})$

$$P(X \leq 14) = P\left(\frac{X-12}{3{,}4} \leq \frac{14-12}{3{,}4}\right) = P(Z \leq 0{,}59) = 0{,}7224$$

La aproximación mejora sustancialmente si utilizamos la corrección por continuidad, calculando $P(X \leq 14{,}5)$ en lugar de $P(X \leq 14)$,

$$P(X \leq 14{,}5) = P\left(\frac{X-12}{3{,}4} \leq \frac{14{,}5-12}{3{,}4}\right) = P(Z \leq 0{,}74) = 0{,}7704$$

La probabilidad de aceptar un lote con $p = 0{,}04$ es 0.77 y la de rechazarlo 0.23. Si se aplica este procedimiento a 200 lotes, el número de lotes rechazados Y es otra vez una variable aleatoria con distribución binomial de parámetros $N = 200$ y $P = 0{,}23$. Entonces el número esperado de lotes rechazados es

$$E[Y] = 200 \times 0{,}23 = 46.$$

• • •

28. Un servicio telefónico de urgencias recibe por término medio 10 llamadas cada minuto, ¿cuál es la probabilidad de recibir más de 550 llamadas en una hora? Se ha diseñado un "*call center*" con capacidad de respuesta de 650 llamadas a la hora, ¿cuál es el número esperado de horas al año con número de llamadas superior a su capacidad? (Se supone que las llamadas son independientes y todas las horas son similares)

Solución:

Si el número de llamadas cada minuto sigue una distribución de Poisson de media 10, el número de llamadas por hora sigue la misma distribución de Poisson de media $\lambda = 600$ llamadas por hora. Por tanto

$$P(Y \geq 551) = P\left(\frac{Y-600}{\sqrt{600}} \geq \frac{551-600}{\sqrt{600}}\right) = P(Z \geq -2{,}00) = 0{,}977$$

La probabilidad de que en una hora haya más de 651 llamadas es:

$$P(Y \geq 651) = P\left(\frac{Y-600}{\sqrt{600}} \geq \frac{651-600}{\sqrt{600}}\right) = P(Z \geq 2{,}08) = 0{,}0188$$

El año tiene $365 \times 24 = 8760$ horas, por tanto el número esperado de horas con demanda superior a 651 es $8760 \times 0{,}0188 = 164{,}7$

• • •

29. A un congreso de medicina acuden 500 personas. Un laboratorio farmaceútico va a regalar corbatas a los hombres y pañuelos a las mujeres. Desgraciadamente no conocen el número exacto de cada sexo, aunque saben de otros congresos que la proporción es similar. Calcula el número mínimo de corbatas y de pañuelos que deben tener disponibles los organizadores para que todos los asistentes tengan el regalo que les corresponde con probabilidad de 0.99 (es decir ninguna mujer se quede sin pañuelo y ningún hombre sin corbata). Se supone que la probabilidad de hombre o mujer es igual a 0.5 y que la probabilidad de que un asistente sea de un determinado sexo es independiente del sexo de los restantes.

Solución:

Llamando X a la variable aleatoria número de hombres que acuden al concierto e Y el número de mujeres. X e Y son variables aleatorias con distribución binomial de parámetros $n = 500$ y $p = 1/2$ que cumplen $X + Y = 500$. Sea a el número de pañuelos y corbatas que se deben comprar para que se cumpla que

$$P(X \le a, Y \le a) = 0{,}99$$

El valor de a es el mismo para pañuelos o corbatas, porque en principio el problema es simétrico para hombres y mujeres. Ahora bien, se cumple que $Y = 500 - X$ por tanto se tiene que:

$$P(X \le a, Y \le a) = P(X \le a, 500 - X \le a) = P(500 - a \le X \le a) = 0{,}99$$

Aproximando X por una normal de media $np = 500 \times 0{,}5 = 250$ y desviación típica $\sqrt{np(1-p)} = \sqrt{500 \times 0{,}5 \times 0{,}5} = \sqrt{125}$, se obtiene que:

$$\begin{aligned} P(500 - a \le X \le a) &= P\left(\frac{(500-a) - 250}{\sqrt{125}} \le \frac{X - 250}{\sqrt{125}} \le \frac{a - 250}{\sqrt{125}}\right) \\ &= P\left(-\frac{a-250}{\sqrt{125}} \le Z \le \frac{a-250}{\sqrt{125}}\right) = 0{,}99 \end{aligned}$$

Z es una variable aleatoria $N(0,1)$, que cumple $P(-2{,}57 \le Z \le 2{,}57) = 0{,}99$ (en las tablas $P(Z \le 2{,}57) = 0{,}995$. Eso significa que:

$$\frac{a - 250}{\sqrt{125}} = 2{,}57, \quad \Rightarrow a = 250 + 2{,}57\sqrt{125} \simeq 279$$

.

Se deben tener disponibles 279 pañuelos y 279 corbatas, para con una probabilidad de 0.99 tener pañuelos suficientes para las mujeres que asistan y corbatas para los hombres.

• • •

30. Federer y Nadal se encuentran empatados, 40-40 en un juego en el que está sacando Nadal. Según las estadísticas la probabilidad de que Nadal gane un punto determinado cuando tiene el saque es 0.6. ¿Cuál es la probabilidad de que el juego lo termine ganando Nadal? (Nota. Piensa en el desempate de la siguiente forma: se juegan dos puntos: si los gana un jugador ese jugador ha ganado el juego, si cada jugador gana un punto se juegan otros dos puntos y se vuelve a aplicar la misma regla).

Solución:

Llamando $p = 0{,}6$ a la probabilidad de que gane Nadal el punto. Para que gane Nadal el juego, debe ganar dos puntos seguidos, la probabilidad de que esto ocurra es $p_N = p^2$. Los dos jugadores empatan si gana un punto cada uno, la probabilidad es $p_E = 2p(1-p)$ (que gane Nadal el primero y pierda el segundo, o al revés, Federer gana el primero y Nadal el segundo). Por último, puede ocurrir que Feder gane los dos puntos seguidos y su probabilidad es $p_F = (1-p)^2$.

Suceso	Resultado	Probabilidad
N	Gana Nadal	p^2
E	Empatan	$2p(1-p)$
F	Gana Federer	$(1-p)^2$

El juego lo gana Nadal si gana los primeros dos puntos o, si empatan en los primeros dos puntos y gana el tercer y cuarto punto o, si empatan en el tercer y cuarto y gana los dos siguientes, y así sucesivamente hasta que se produzca el desempate.

$$P(GanaNadal) = p_N + p_E \times p_N + p_E^2 \times p_N + p_E^3 \times p_N + \cdots = \frac{p_N}{1-p_E} = \frac{p^2}{1-2p(1-p)}$$

La serie anterior es la suma de una progresión geométrica de razón p_E. Se puede comprobar que $p^2 + (1-p)^2 = 1 - 2p(1-p)$, otra forma de escribir el resultado anterior es:

$$P(GanaNadal) = \frac{p^2}{p^2 + (1-p)^2} = \frac{0{,}6^2}{0{,}6^2 + 0{,}4^2} = 0{,}692$$

• • •

31. *"Cibeles in Concert"* es una empresa que organiza viajes en autobus para asistir a actuaciones musicales. Para un concierto de *Bruce Springsteen* en Paris ha ofertado 300 plazas que salen de Madrid y Sevilla. Las reservas se hacen por Internet, el precio del viaje para los de Sevilla es de 60 euros y para los de Madrid 50 euros . Las 300 plazas se cubren con seguridad. Si la probabilidad de que un asistente salga de Sevilla es 1/3 y de Madrid 2/3, y se acepta independencia entre las 300 reservas, calcula los ingresos esperados por la compañía y la varianza de estos ingresos.

Solución:

El número de clientes X que parten de Sevilla es una variable aleatoria binomial de parámetros $n = 300$ y $p = 1/3$, el número de participantes Y que salen de Madrid es una variable binomial de parámetros $n = 300$ y $p = 2/3$. Según el enunciado $X + Y = 300$. Los ingresos de los organizadores del viaje son:

$$\begin{aligned} U &= 60X + 50Y \\ &= 60X + 50(300 - X) \\ &= 15000 + 10X \end{aligned} \tag{3.1}$$

Como $E[X] = 300 \times 1/3 = 100$ y $Var(X) = 300 \times \dfrac{1}{3} \times \dfrac{2}{3}$, entonces

$$\begin{aligned} E[U] &= 15000 + 10E[X] = 16000 \\ Var[U] &= 10^2 Var[X] = 10^2 \times 300 \times \frac{1}{3} \times \frac{2}{3} = 6666{,}6 \end{aligned} \tag{3.2}$$

Aproximando la distribución binomial por la normal, los ingresos tienen distribición normal de media 16000 y desviación típica 816.

• • •

32. Una empresa de celulosa tiene dos líneas para fabricar pasta de papel en planchas de 1m × 1m. Una medida de su calidad es la limpieza, que se mide en número de impurezas (partículas) por m^2. La línea I, fabrica con una tasa media de 5 impurezas por m^2 y la línea II con 3 impurezas por m^2. El número de impurezas por plancha es una variable aleatoria con distribución de Poisson. Las planchas de pasta se empaquetan en balas de 2000 unidades y se almacenan clasificadas según la línea de procedencia. Cuando por algún motivo se encuentra en el almacén una bala sin clasificar se adopta el siguiente criterio: tomar una muestra aleatoria de 10 planchas y determinar el número medio de impurezas. Asignarla al grupo de la línea I si el número medio de impurezas es mayor que 4 y a la línea II en caso contrario. (Se supone que la probabilidad inicial de pertenecer a una u otra linea es la misma).

   a) Calcular la probabilidad de clasificar erróneamente una bala.

   b) En un caso concreto, el número de impurezas en cada una de las diez planchas han sido $1, 2, 3, 3, 4, 4, 5, 5, 6, 8$. Calcular la probabilidad de que pertenezca a cada una de las líneas.

Solución:

a) Según el enunciado X_i el número de impurezas encontrados en la plancha i es una variable aleatoria con distribución de Poisson. Llamando λ al parámetro de la distribución, se tiene que para una muestra $X_1, X_2, ..., X_{10}$, la media

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_{10}}{10}$$

cumple

$$E[\bar{X}] = \lambda \quad Var[\bar{X}] = \frac{\lambda}{10}$$

La distribución de probabilidad de $\bar{X}$ se puede aproximar por una Normal. Si la bala es de la línea I, $\bar{X} \sim N(5, \sqrt{\frac{5}{10}})$ y si procede de la línea II, $\bar{X} \sim N(3, \sqrt{\frac{3}{10}})$.

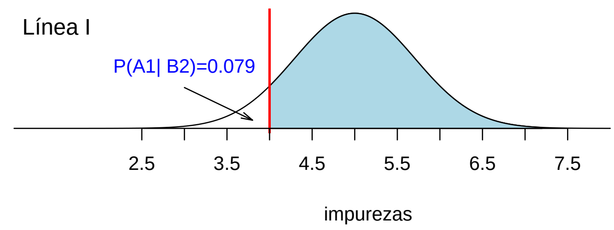
	A_1	A_2
$B_1 = \{\bar{X} > 4\}$	Bien	Error
$B_2 = \{\bar{X} \leq 4\}$	Error	Bien

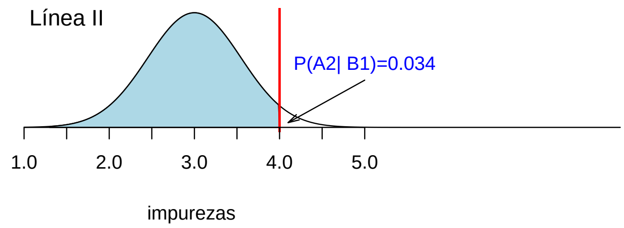
Sea A_1 el suceso la bala pertenece a la línea I y A_2 la bala pertenece a la línea II, B_1 la bala se clasifica en la línea I y B_2 la bala se clasifica en la línea II. Llamando E al suceso error, está definido por $E = (A_1 \cap B_2) \cup (A_2 \cap B_1)$

$$\begin{aligned}
P(E) &= P\left[(A_1 \cap B_2) \cup (A_2 \cap B_1)\right] \\
&= P(B_2|A_1)P(A_1) + P(B_1|A_2)P(A_2) \\
&= P(\bar{X} \leq 4|A_1)P(A_1) + P(\bar{X} > 4|A_2)P(A_2) \\
&= P\left(\frac{\bar{X}-5}{\sqrt{5/10}} \leq \frac{4-5}{\sqrt{5/10}}\right) \times \frac{1}{2} + P\left(\frac{\bar{X}-3}{\sqrt{3/10}} \geq \frac{4-3}{\sqrt{3/10}}\right) \times \frac{1}{2} \\
&= P(Z \leq -1{,}41) \times \frac{1}{2} + P(Z \geq 1{,}82) \times \frac{1}{2} \\
&= 0{,}0563
\end{aligned}$$

b) Llamando $C = \{X_1 = 1, X_2 = 2, ..., X_{10} = 8\}$,

$$\begin{aligned}
P(C|A_1) &= e^{-5}\frac{5^1}{1!} \times e^{-5}\frac{5^2}{2!} \cdots e^{-5}\frac{5^8}{8!} \\
&= e^{-50}\frac{5^{41}}{1!2!3! \cdots 8!} \\
P(C|A_2) &= e^{-30}\frac{3^{41}}{1!2!3! \cdots 8!}
\end{aligned} \tag{3.3}$$





y

$$\begin{aligned} P(C) &= P(C|A_1)P(A_1) + P(C|A_2)P(A_2) \\ &= e^{-50}\frac{5^{41}}{1!2!3!\cdots 8!} \times \frac{1}{2} + e^{-30}\frac{3^{41}}{1!2!3!\cdots 8!} \times \frac{1}{2} \end{aligned} \tag{3.4}$$

y piden la probabilidad de que venga de la línea I,

$$\begin{aligned} P(A_1|C) &= \frac{P(C|A_1)P(A_1)}{P(C)} \\ &= \frac{e^{-50}\frac{5^{41}}{1!2!3!\cdots 8!}}{e^{-50}\frac{5^{41}}{1!2!3!\cdots 8!} + e^{-30}\frac{3^{41}}{1!2!3!\cdots 8!}} \\ &= \frac{e^{-50} \times 5^{41}}{e^{-50} \times 5^{41} + e^{-30} \times 3^{41}} = 0{,}7199 \end{aligned} \tag{3.5}$$

La probabilidad de que provenga de la línea II es $P(A_2|C) = 1 - P(A_1|C) = 0{,}2801$

• • •

33. Los billetes de banco son fabricados en pliegos. La impresión se realiza por dos máquinas iguales, una de ellas imprime el anverso y la otra el reverso. Sea X e Y, respectivamente, el número de defectos de impresión en el anverso y reverso de un pliego. Ambas variables son independientes con distribución de Poisson de parámetros λ_1 y λ_2.

a) Demostrar que el número total de defectos en un pliego $Z = X + Y$ tiene distribución de Poisson. (Nota.- Utilizar que

$$\mathrm{P}\{Z = n\} = \sum_{k=0}^{n} \mathrm{P}\{X = k\}\mathrm{P}\{Y = n - k\}$$

y el desarrollo del binomio de Newton para $(\lambda_1 + \lambda_2)^n$.)

b) Si el número total de defectos en un pliego es $Z = n$, ¿ cuál es la probabilidad de que haya exáctamente $X = k$ defectos en el anverso? (Obtener la expresión en función de λ_1, λ_2, n y k). ¿ De qué distribución de probabilidad se trata?

Solución:

a)

$$\begin{aligned}
\mathrm{P}\{Z = n\} &= \sum_{k=0}^{n} \mathrm{P}\{X = k, Y = n - k\} && \text{(convolución)} \\
&= \sum_{k=0}^{n} \mathrm{P}\{X = k\}\mathrm{P}\{Y = n - k\} && \text{(independencia)} \\
&= \sum_{k=0}^{n} \frac{e^{-\lambda_1}\lambda_1^k}{k!} \times \frac{e^{-\lambda_2}\lambda_2^{(n-k)}}{(n-k)!} && \text{(Poisson)} \\
&= e^{-(\lambda_1+\lambda_2)} \sum_{k=0}^{n} \frac{\lambda_1^k}{k!} \times \frac{\lambda_2^{(n-k)}}{(n-k)!} && \text{(saco factor común } e^{-(\lambda_1+\lambda_2)}\text{)} \\
&= e^{-(\lambda_1+\lambda_2)} \sum_{k=0}^{n} \frac{1}{k!(n-k)!} \times \lambda_1^k \lambda_2^{(n-k)} && \text{(reordeno)} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{n!} \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} \times \lambda_1^k \lambda_2^{(n-k)} && \text{(multiplico y divido por n!)} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{n!}(\lambda_1 + \lambda_2)^n && \text{(binomio de Newton)}
\end{aligned} \tag{3.6}$$

Se comprueba que Z tiene distribución de Poisson de parámetro $\lambda_1 + \lambda_2$.

b) Se trata de calcular la probabilidad condicionada $P(X = k|Z = n)$:

$$\begin{aligned}
P(X=k|Z=n) &= \frac{P(X=k,Z=n)}{P(Z=n)} \\
&= \frac{P(X=k,Y=n-k)}{P(Z=n)} \\
&= \frac{P(X=k)P(Y=n-k)}{P(Z=n)} \\
&= \frac{\frac{e^{-\lambda_1}\lambda_1^k}{k!} \times \frac{e^{-\lambda_2}\lambda_2^{(n-k)}}{(n-k)!}}{\frac{e^{-(\lambda_1+\lambda_2)}}{n!}(\lambda_1+\lambda_2)^n} \\
&= \frac{n!}{k!\,(n-k)!}\frac{\lambda_1^k\lambda_2^{n-k}}{(\lambda_1+\lambda_2)^n} \\
&= \binom{n}{k}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{n-k} \\
&= \binom{n}{k}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k\left(1-\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^{n-k}
\end{aligned}$$

que es la distribución binomial de parámetros n y $p = \dfrac{\lambda_1}{\lambda_1+\lambda_2}$.

• • •

34. La llegada de los clientes a un banco se considera un proceso Poisson con parámetro λ. Sabiendo que en la última hora han llegado 2 clientes, ¿cuál es la probabilidad de que los dos entraran en los primeros 15 minutos?

Solución:

Dividimos el número de clientes que llegan en una hora Z en dos variables: la variable aleatoria X *" número de clientes en los primeros 15 min"*, e Y, *" número de clientes en los siguientes 45 min"*. X e Y son variables aleatorias independientes con distribuci ón de Poisson de parámetros $\lambda/4$ y $3\lambda/4$, respectivamente. Aplicamos el mismo razonamiento que en el ejercicio 33,

$$\begin{aligned}
P(X=2|Z=2) &= \frac{P(X=2,Z=2)}{P(Z=2)} \\
&= \frac{P(X=2,Y=0)}{P(Z=2)} \\
&= \frac{P(X=2)P(Y=0)}{P(Z=2)} \\
&= \left(\frac{e^{-\lambda/4}\,(\lambda/4)^2}{2!} \times \frac{e^{-3\lambda/4}\,(3\lambda/4)^0}{0!}\right) / \frac{e^{-\lambda}}{2!}\lambda^2 \\
&= \left(\frac{1}{4}\right)^2
\end{aligned}$$

Como se ha mencionada, este es una caso particular del resultado obtenido en el ejercicio 33, si conocemos el número total n de clientes que han acudido en una hora, el número de clientes que entraron en el primer cuarto de hora es una variable aleatoria con distribución binomial de parámetros n y $p = 1/4$. En este caso $n = 2$ y

$$P(X=2|Z=2) = \binom{2}{2}\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^0.$$

• • •

35. Un equipo de radio tiene dos partes, el receptor y el amplificador. La duración del receptor es una variable aleatoria exponencial de media 500 horas y la duración del amplificador una variable exponencial de media 1000 horas. ¿Cuál es la probabilidad de que el fallo del equipo (cuando se produzca) sea debido a un fallo del receptor? (Se supone que las variables son independientes)

Solución:

Llamando X e Y a las variables aleatorias tiempo de vida del receptor y del emisor, se pide calcular la $P(X > Y)$. Para ello se debe calcular

$$P(X > Y) = \iint\limits_{x>y} f_{X,Y}(x,y)\mathrm{d}x\,\mathrm{d}y$$

siendo $f_{X,Y}$ la función de densidad conjunta de las dos variables.

Según el enunciado $f_X(x) = ae^{-ax}$, $x > 0$ y $f_Y(y) = be^{-by}$, $y > 0$ y como son independientes la función de densidad conjunta es:

$$f_{X,Y}(x,y) = a\,be^{-ax-by}, \quad x > 0, y > 0,$$

con $a = 1/500$ y $b = 1/1000$.

$$
\begin{aligned}
P(X > Y) &= \int_0^{\infty} \int_y^{\infty} a\, b e^{-ax-by} \mathrm{d}x\, \mathrm{d}y \\
&= \int_0^{\infty} a\, b e^{-by} \left[\int_y^{\infty} e^{-ax} \mathrm{d}x \right] \mathrm{d}y \\
&= \int_0^{\infty} a\, b e^{-by} \left[-\frac{1}{a} e^{-ax} \right]_y^{\infty} \mathrm{d}y \\
&= \int_0^{\infty} b\, e^{-(a+b)y}\, \mathrm{d}y \\
&= \left[-\frac{b}{a+b} e^{-(a+b)y} \right]_0^{\infty} \\
&= \frac{b}{a+b} = \frac{1/1000}{1/500 \, + 1/1000} \\
&= \frac{500}{500 \, + \, 1000} = \frac{1}{3}.
\end{aligned}
$$

• • •

36. Sea X_1 una variable aleatoria N(10,1), X_2 una variable aleatoria N(20,1), y X_3 una variable aleatoria N(30,4). Se define

$$Z_1 = X_1 + X_2 - X_3$$

$$Z_2 = X_1 + X_2 + X_3$$

$$Z_3 = X_1 - X_2 - X_3$$

Si X_1, X_2, X_3 son independientes, calcular la matriz de varianzas de (Z_1, Z_2, Z_3).

Solución:

Escribiendo las ecuaciones anteriores de forma matricial,

$$
\begin{pmatrix} Z_1 \\ Z_2 \\ Z_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \\ 1 & -1 & -1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \tag{3.7}
$$

o lo que es lo mismo $\mathbf{Z} = \mathbf{AX}$ donde

$$
\mathbf{Z} = \begin{pmatrix} Z_1 \\ Z_2 \\ Z_3 \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \\ 1 & -1 & -1 \end{pmatrix} \quad \text{y} \quad \mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}
$$

La matriz de varianzas de $\mathbf{Z}$ es igual a

$$\begin{aligned} Var[\mathbf{Z}] &= \mathbf{A}Var[\mathbf{X}]\mathbf{A}^T \\ &= \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \\ 1 & -1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 16 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ -1 & 1 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 18 & -14 & 16 \\ -14 & 18 & -16 \\ 16 & -16 & 18 \end{pmatrix}. \end{aligned}$$

• • •

37. Una oficina de correos tiene dos ventanillas de atención al público. Tres personas A,B y C llegan en el mismo instante a la oficina de correos y encuentran las dos ventanillas desocupadas. Los tiempos de servicio requeridos por las tres personas son variables aleatorias independientes con distribución exponencial de parámetro λ. Los tiempos de servicio de A y B comienzan de inmediato, mientras que C debe esperar a que termine el primero de los dos. ¿Cuál es la probabilidad de que C no sea el último en salir de la oficina de correos?

Solución:

A y B tienen la misma probabilidad de acabar primero. Sea X, Y y Z los tiempos que están cada uno en la oficina de correos. Supongamos que termina A primero en el instante a, la probabilidad de que B necesite t horas más es

$$P(Y > a+t|Y > a) = \frac{P(Y > a+t, Y > a)}{P(Y > a)} = \frac{P(Y > t+a)}{P(Y > a)} = \frac{e^{-\lambda(t+a)}}{e^{-\lambda a}} = e^{-\lambda t}$$

C empieza justo en el instante a y la probabilidad de que necesite más de t horas es $P(Z > t) = e^{-\lambda t}$. En el instante en que termina A, los otros dos B y C están en las mismas condiciones y por tanto, tienen las misma probabilidad de acabar segundo. El orden de salida puede ser cualquieras de los cuatro que se muestran en la tabla 3.3 Todas las combinaciones tienen la misma probabilidad. La probabilidad

Tabla 3.3: Orden de Salida

1	2	3	Probabilidad
A	B	C	1/4
A	C	B	1/4
B	A	C	1/4
B	C	A	1/4

de que C sea el último es 1/2. A y B tienen la misma probabilidad igual 1/4 de ser el último.

• • •

38. En cierta fabricación mecánica el 96 % de las piezas resultan con longitudes admisibles (dentro de tolerancias), un 3 % son piezas defectuosas cortas y un 1 % son defectuosas largas. Calcular la probabilidad de:

    a) En un lote de 250 piezas sean admisibles 242 o más.

    b) En un lote de 500 sean cortas 10 o menos.

    c) En 1000 piezas haya entre 6 y 12 largas. Todas las aproximaciones se calculan la distribución normal.

    Solución:

    a) El número de piezas admisibles sigue la distribución binomial de parámetros $n = 250$ y $p = 0{,}96$, aproximando por la normal $X \sim N(np, \sqrt{np(1-p)})$, siendo la media $np = 240$ y la desviación típica $\sqrt{np(1-p} = 3{,}1$:

    $$\begin{aligned} P(X \geq 242{,}5) &= P\left(\frac{X-240}{3{,}1} \geq \frac{242{,}5-240}{3{,}1}\right) \\ &= P(Z \geq 0{,}8235) = 0{,}2098 \end{aligned}$$

    b) El número de piezas cortas Y en 500 se puede aproximar por una normal de media $500 \times 0{,}03 = 15$ y desviación típica $\sqrt{500 \times 0{,}97 \times 0{,}03} = 3{,}81$,

    $$\begin{aligned} P(Y \leq 10{,}5) &= P\left(\frac{X-15}{3{,}81} \leq \frac{10{,}5-15}{3{,}81}\right) \\ &= P(Z \leq -1{,}18) = 0{,}1191 \end{aligned}$$

    c) El número de piezas largas U en 1000 se puede aproximar por una normal de media $1000 \times 0{,}01 = 10$ y desviación típica $\sqrt{1000 \times 0{,}01 \times 0{,}99} = 3{,}15$,

    $$\begin{aligned} P(5{,}5 \leq U \leq 12{,}5) &= P\left(\frac{5{,}5-10}{3{,}15} \leq \frac{U-10}{3{,}15} \leq \frac{12{,}5-10}{3{,}15}\right) \\ &= P(-1{,}43 \leq Z \leq 0{,}7945) \\ &= P(Z \leq 0{,}7945) - P(Z \leq -1{,}43) \\ &= 0{,}7865 - 0{,}0764 \\ &= 0{,}7101 \end{aligned}$$

• • •

39. La estatura de los ciudadanos varones de un país sigue una distribución normal:

    $$N(\mu = 175cm, \sigma = 5cm);$$

si se seleccionan al azar 100 ciudadanos, ¿cuál es la probabilidad de que al menos 30 superen los 180cm?

Solución:

La probabilidad de que un varón mida más de 180 cm es

$$\begin{aligned} P(X \geq 180) &= P\left(\frac{X-175}{5} \geq \frac{180-175}{5}\right) \\ &= P(Z \geq 1) = 0{,}1587 \end{aligned}$$

El número de varones que miden más de 180 en la muestra de 100 sigue una distribución binomial con $n = 100$ y $p = 0{,}1587$, se aproxima por una normal de media $np = 15{,}87$ y desviación típica $\sqrt{np(1-p} = 3{,}65$. La probabilidad de que haya más de 30 es (utilizando la correción por continuidad):

$$\begin{aligned} P(X \geq 29{,}5) &= P\left(\frac{X-15{,}87}{3{,}65} \geq \frac{29{,}5-15{,}87}{3{,}65}\right) \\ &= P(Z \geq 3{,}732) = 0{,}0001 \end{aligned}$$

• • •

40. En un proceso de fabricación de película fotográfica aparecen por término medio 1 defecto de cada 20 metros de película. Si la distribución de defectos es Poisson, calcular la probabilidad de 6 defectos en un rollo de 200 metros de pelicula (a) directamente; (b) utilizando la aproximación normal.

Solución:

El número de defectos en un rollo de 200 m sigue una distribución de Poisson de parámetros $\lambda = 10$, la probabilidad de exáctamente 6 defectos es

$$P(X = 6) = e^{-10}\frac{10^6}{6!} = 0{,}063$$

si utilizamos la aproximación normal:

$$\begin{aligned} P(5{,}5 \leq X \leq 6{,}5) &= P\left(\frac{5{,}5-10}{\sqrt{10}} \leq \frac{X-10}{\sqrt{10}} \leq \frac{6{,}5-10}{\sqrt{10}}\right) \\ &= P(-1{,}423 \leq Z \leq 1{,}107) \\ &= P(Z \leq 1{,}107) - P(Z \leq -1{,}423) \\ &= 0{,}1344 - 0{,}0774 \\ &= 0{,}0570 \end{aligned}$$

• • •

41. De un lote que contiene 200.000 piezas se extraen al azar 250. Si el número de piezas defectuosas en la muestra es mayor que c se rechaza el lote y se devuelve al proveedor. Calcular c si se desea que la probabilidad de rechazar un lote con 3 % de piezas defectuosas sea 0.05. (Nota: suponer que se toma la muestra con remplazamiento para hacer los cálculos y utilizar la aproximación normal para obtener c).

    Solución: El número de defectuosos en la muestra X sigue una distribución Binomial con parámetros $n = 250$ y $p = 0{,}03$. Se desea calcular c tal que $P(X > c) = 0{,}05$, es decir

    $$P(X \leq c) = \sum_{k=0}^{c} \binom{250}{k} 0{,}03^k 0{,}97^{250-k} = 0{,}95$$

    Utilizando en `R` la instrucción `qbinom(0.95,size=250,prob=0.03)`, el resultado es $c = 12$. Empleando la aproximación de la Binomial por la Normal, se tiene que X se puede aproximar a una Normal de media $np = 7{,}5$ y desviación típica igual a $\sqrt{np(1-p)} = 2{,}7$, de manera que:

    $$\begin{aligned} P(X \leq c) &= P(\frac{X - 7{,}5}{2{,}7} \leq \frac{c - 7{,}5}{2{,}7}) \\ &= P(Z \leq \frac{c - 7{,}5}{2{,}7}) = 0{,}95 \end{aligned}$$

    Las tablas de la Normal (0,1), $P(Z \leq 1{,}645) = 0{,}95$, por tanto

    $$\frac{c - 7{,}5}{2{,}7} = 1{,}645 \Rightarrow c = 7{,}5 + 1{,}645 \times 2{,}7 = 11{,}94 \approx 12.$$

• • •

42. Para controlar la recepción de lotes de 10000 unidades de discos compactos se toma una muestra al azar de $n = 200$ discos clasificándolos como aceptables y defectuosos. Si el número de discos defectuosos es igual o inferior a $c = 15$ se acepta el lote, en caso contrario se rechaza. ¿Cuál es la probabilidad de aceptar un lote con el 12 % de discos defectuosos? ¿Cuál es la probabilidad de rechazar un lote con el 5 % de discos defectuosos? ¿Que valores de n y c deben utilizarse si se desea que las probabilidades anteriores sean iguales a 0.05?

    Solución:

    X la variable aleatoria *número de discos defectuosos* tiene distribución binomial. Si el lote tiene 12 % de piezas defectuosas, la probabilidad de aceptarlo es

    $$P(X \leq 15) = \sum_{x=0}^{15} \binom{200}{x} 0{,}12^x 0{,}88^{100-x} = 0{,}0267$$

Utilizando la aproximación normal, $X \sim N(np, \sqrt{np(1-p)})$, con $np = 24$ y desviación típica

$$\sqrt{np(1-p)} = 4{,}6,$$

y utilizando la corrección por continuidad, se tiene que:

$$P(X \leq 15{,}5) = P\left(\frac{X-24}{4{,}6} \leq \frac{15{,}5-24}{4{,}6}\right) = P(Z < -1{,}85) = 0{,}0321$$

La probabilidad de aceptar un lote con el 5 % de defectuosos es

$$P(X \leq 15) = \sum_{x=0}^{15} \binom{200}{x} 0{,}05^x 0{,}95^{100-x} = 0{,}9556$$

y utilizando la aproximación normal, con media $np = 10$ y desviación típica

$$\sqrt{np(1-p)} = 3{,}08,$$

se tiene que

$$P(X \leq 15{,}5) = P\left(\frac{X-10}{3{,}08} \leq \frac{15{,}5-10}{3{,}08}\right) = P(Z < 1{,}784) = 0{,}9628$$

Por este método, la probabilidad de rechazarlo es $P(X \geq 15{,}5) = 0{,}0372$

Se desea calcular n (el tamaño muestral) y c (el número máximo de piezas defectuosas de la muestra para que el lote ses aceptado), de manera que la probabilidad de rechazar un lote con 5 % defectuosas sea 0.05 y la probabilidad de aceptar un lote con el 12 % de piezas defectuosas sea 0,05.

Este problema es clásico en control de recepción de lotes de piezas. El planteamiento consiste en: en primer lugar, decidir lo que se considera un lote bueno o aceptable para el comprador, por ejemplo, aquel que contenga el $p_a = 5\,\%$ de piezas defectuosas. Y lo que es un lote inaceptable o "malo"para el comprador, por ejemplo si contiene el $p_r = 12\,\%$ de piezas defectuosas. Estos valores se eligen de común acuerdo entre el comprador y vendedor. Tienen que ser valores razonablemente distintos, usualmente p_r es dos o tres veces p_a. En segundo lugar se establecen los riesgos: probabilidad de aceptar un lote malo (riesgo del comprador) y probabilidad de rechazar un lote bueno (riesgo del vendedor). En este ejercicio el riesgo del vendedor y comprador es el mismo, igual a 0.05.

El problema se plantea como un sistema con dos ecuaciones y dos incógnitas n y c, la primera ecuación se obtiene de la condición $P(X \leq c|p = 0{,}05) = 0{,}95$ y la segunda ecuación de $P(X \leq c|p = 0{,}12) = 0{,}05$.

La primera ecuación es la condición que pone el *vendedor*, que si el lote es "bueno" $p = 0{,}05$ sea aceptado con alta probabilidad (0.95). Utilizando la aproximación normal se tiene:

$$P(X \leq c|p = 0{,}05) = P\left(\frac{X - n \times 0{,}05}{\sqrt{n \times 0{,}05 \times (1 - 0{,}05)}} \leq \frac{c - n \times 0{,}05}{\sqrt{n \times 0{,}05 \times (1 - 0{,}05)}}\right) = 0{,}95$$

En las tablas de la normal se obtiene el valor $z_a = 1{,}645$ que cumple $P(Z \leq z_a) = 0{,}95$, de manera que la ecuación del vendedor queda:

$$\frac{c - n \times 0{,}05}{\sqrt{n \times 0{,}05 \times (1 - 0{,}05)}} = 1{,}645 \tag{3.8}$$

La segunda ecuación es la condición que pone el *comprador*, que si el lote es "malo" $p = 0{,}12$ sea aceptado con baja probabilidad (0.05):

$$P(X \leq c|p = 0{,}12) = P\left(\frac{X - n \times 0{,}12}{\sqrt{n \times 0{,}12 \times (1 - 0{,}12)}} \leq \frac{c - n \times 0{,}12}{\sqrt{n \times 0{,}12 \times (1 - 0{,}12)}}\right) = 0{,}05$$

En las tablas de la normal se obtiene el valor $z_r = -1{,}645$ que cumple $P(Z \leq z_r) = 0{,}05$, de manera que la ecuación del comprador queda:

$$\frac{c - n \times 0{,}12}{\sqrt{n \times 0{,}12 \times (1 - 0{,}12)}} = -1{,}645 \tag{3.9}$$

Uniendo las ecuaciones del vendedor (3.8) y la del comprador (3.9) se forma el sistema

$$\begin{aligned} c &= n \times 0{,}05 + 1{,}645\sqrt{n \times 0{,}05 \times 0{,}95} \\ c &= n \times 0{,}12 - 1{,}645\sqrt{n \times 0{,}12 \times 0{,}88} \end{aligned}$$

Igualando las dos ecuaciones y despejando n se tiene

$$n = \left(\frac{1{,}645\sqrt{0{,}05 \times 0{,}95} + 1{,}645\sqrt{0{,}12 \times 0{,}88}}{0{,}12 - 0{,}05}\right)^2 = 163$$

y sustituyendo en la primera ecuación del sistema se obtiene c

$$c = 163 \times 0{,}05 + 1{,}645\sqrt{163 \times 0{,}05 \times 0{,}95} = 13$$

• • •

43. La función de densidad conjunta de las variables aleatorias X, Y es la siguiente normal bidimensional:

$$f(x,y) = \frac{1}{2\pi(1-\rho^2)^{1/2}} \exp\left(\frac{-1}{2(1-\rho^2)}(x^2 - 2\rho xy + y^2)\right)$$

Si $\rho = 0{,}3$, calcular $\Pr(Y \leq X+1)$.

Solución:

La función de densidad anterior corresponde a la normal bi-dimensional. La inspección de la ecuación indica que las dos variables aleatoris X e Y tienen media cero y varianza unidad. La covarianza por tanto es igual a 0.3. Teniendo en cuenta todo ello:

$$P(Y \leq X+1) = P(Y - X \leq 1)$$

La variable $U = Y - X$ tiene distribución normal de media 0 y varianza 1.4:

$$Var(Y-X) = Var(Y) + Var(X) - 2Cov(X,Y) = 1 + 1 - 2 \times 0{,}3 = 1{,}4$$

$$P(Y \leq X+1) = P(U \leq 1) = P(\frac{U}{\sqrt{1{,}4}} \leq \frac{1}{\sqrt{1{,}4}})$$

$$P(Y \leq X+1) = P(Z \leq 0{,}845) = 0{,}800$$

• • •

44. Un inversor tiene su dinero repartido en acciones de dos compañías. La rentabilidad (% de beneficio) de las compañías pueden ser consideradas como dos variables aleatorias, con media anual igual para las dos del 10 %, aunque el riesgo es muy diferente. Una tiene una desviación típica de 2.5 % y la otra del 1 %. Además se sabe que la correlación entre ellas es -0.50. ¿Qué proporción debe invertir en cada una para que el riesgo sea mínimo? (Nota: Mínimo riesgo es lo mismo que mínima varianza. $X \rightarrow N(10, 2{,}5)$ y $Y \rightarrow N(10, 1)$, $corr(X, Y) = -0{,}5$).

Solución:

Sea α la proporción invertida en la primera compañía y $1-\alpha$ la proporción invertida en la segunda. La rentabilidad total será $U = \alpha X + (1-\alpha)Y$. La varianza de U se obtiene como:

$$Var(U) = \alpha^2\sigma_X^2 + (1-\alpha)^2\sigma_Y^2 + 2\alpha(1-\alpha)\sigma_{XY}$$

El valor α que minimiza la $Var(U)$ se obtiene igualando la derivada a cero:

$$\frac{dVar(U)}{d\alpha} = 2\alpha\sigma_X^2 - 2(1-\alpha)\sigma_Y^2 + 2(1-2\alpha)\sigma_{XY}$$

Igualando a cero y despejando α se tiene:

$$\alpha = \frac{\sigma_Y^2 - \sigma_{XY}}{\sigma_X^2 + \sigma_Y^2 - 2\sigma_{XY}}$$

$$\alpha = \frac{1 + 0{,}5 \times 2{,}5}{2{,}5^2 + 1^2 + 2 \times 0{,}5 \times 2{,}5} = 0{,}23$$

Se ha utilizado $\sigma_{XY} = \rho\sigma_X\sigma_Y = -0{,}5 \times 1 \times 2{,}5$. Haciendo la segunda derivada se tiene que

$$\frac{d^2Var(U)}{d\alpha^2} = 2\sigma_X^2 + 2\sigma_Y^2 - 4\sigma_{XY} = 2(2{,}5^2 + 1 + 2 \times 0{,}5 \times 1 \times 2{,}5) > 0$$

por tanto para $\alpha = 0{,}23$ se tiene una combinación con rentabilidad del 10 y varianza mínima igual a 0.48. (Este valor se obtiene sustituyendo $\alpha = 0{,}23$ en la fórmula $Var(U)$.

• • •

45. Cierto compuesto metálico se ha sometido durante una hora a una atmósfera de oxígeno a 200 grados centígrados. Una medida de su corrosión es la ganancia de peso durante este tiempo. Se ha comprobado que para un determinado compuesto esta ganancia X_1, en una hora, se distribuye segúun una normal $N(100, 5)$.

   a) Si se realizan los ensayos de manera secuencial, ¿cuántos se tendrán que hacer por término medio para encontrar una probeta con una ganancia mayor que 105?

   b) Se comprueba que sometiendo la probeta al ensayo durante dos horas, la ganancia en la segunda hora X_2 tiene una distribución normal $N(60, 5)$, siendo el coeficiente de correlación entre las variables X_1 y X_2, $\rho = -0{,}28$. Calcular la probabilidad de que una probeta tenga mayor ganancia de peso en la segunda hora que en la primera.

   c) Se forman los índices de oxidación $Z_1 = X_1 + X_2$ y $Z_1 = X_1 - X_2$. Calcular la función de densidad conjunta de estas dos nuevas variables. ¿Son independientes?

   d) Si la ganancia durante las dos horas ha sido 170, ¿cuál es el valor medio de la ganancia en la primera hora?

Solución:

a)

$$P(X_1 > 105) = P(Z > \frac{105 - 100}{5}) = P(Z > 1) = 0{,}1586$$

El número Y de ensayos que hay que hacer hasta que aparece la primera probeta con ganancia mayor de 105 es una variable aleatoria *geométrica* de parámetro $p = 0{,}1586$ que tiene como media:

$$E[Y] = \frac{1}{p} = 6{,}3$$

b) La covarianza entre X_1 y X_2 es igual a $-02{,}8 \times 5 \times 5 = -7$, la distribución conjunta de $(X_1, X_2)^T$ es

$$\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N\left[\begin{pmatrix} 100 \\ 60 \end{pmatrix}, \begin{pmatrix} 25 & -7 \\ -7 & 25 \end{pmatrix}\right]$$

Y se desea calcular $P(X_2 > X_1) = P(X_2 - X_1 > 0)$, llamando $U = X_2 - X_1$, se tiene que

$$E[U] = E[X_2] - E[X_1] = 60 - 100 = -40$$

$$Var[U] = Var[X_2] + Var[X_2] - 2Cov(X_1, X_2) = 25 + 25 + 14 = 64$$

y $P(U > 0) = P(Z > (0 + 40)/8) = P(Z > 5) = 0$, siendo $Z \sim N(0, 1)$.

c) El nuevo vector de variables aleatorias $(Z_1, Z_2)^T$ definido como

$$\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$$

tiene distribución normal bivariada con la siguiente media y matriz de varianzas

$$E\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 100 \\ 60 \end{pmatrix} = \begin{pmatrix} 160 \\ 40 \end{pmatrix}$$

$$Var\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 25 & -7 \\ -7 & 25 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 36 & 0 \\ 0 & 64 \end{pmatrix}.$$

d) Dadas dos variables aleatorias X e Y con distribución conjunta normal bivariante

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N\left[\begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, \begin{pmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{pmatrix}\right]$$

se puede demostrar que la distribución de Y condicionada a $X = x$ es normal con media igual a

$$E[Y|X = x] = \mu_Y + \frac{\sigma_{XY}}{\sigma_X^2}(x - \mu_X)$$

$$Var[Y|X = x] = (1 - \rho_{XY}^2)\sigma_Y^2$$

En este caso nos piden la distribución de X_1 condicionada a que $Z_1 = 170$, teniendo en cuenta que

$$\mathrm{Cov}(X1, Z_1) = \mathrm{Cov}(X_1, X_1 + X_2) = \mathrm{Cov}(X_1, X_1) + \mathrm{Cov}(X_1, X_2) = 25 - 7 = 18$$

$$Cor(X_1, Z_1) = 18/(5 \times 6) = 0{,}6$$

se deduce que

$$E[X_1|Z_1 = 170] = 100 + \frac{18}{36}(170 - 160) = 105$$

$$\mathrm{Var}[X_1|Z_1 = 170] = 25(1 - 0{,}6^2) = 16.$$

• • •

46. El abastecimiento de energía eléctrica de una comarca depende de tres centrales: una nuclear, una térmica de carbón y una hidráulica con potencias instaladas de 500 MW, 300 MW y 200 MW, respectivamente. Desde el punto de vista de fiabilidad, cada central sólo puede estar en uno de estos dos estados: disponible (con toda su potencia) o averiada (con potencia cero). En un día, la probabilidad de avería de la central nuclear es 0,10, de la térmica es 0,12 y de la hidráulica 0,05. Las averías son independientes y supondremos como hipótesis simplificadora que a lo largo de un día la central no cambia de estado.

    a) Para un día, calcular la función de distribución de la variable aleatoria $X =$ "Potencia disponible en la comarca".

    b) Un pais tiene 30 centrales de cada uno de los tipos del apartado 1. Calcular la probabilidad de que en un día la potencia disponible en el pais sea menor que 24000 MW. (Utilizar la aproximación normal).

    c) La potencia máxima diaria demandada en el pais es una variable aleatoria con distribución normal de media 23000 MW y desviación típica 1000 MW. ¿Cuál es la probabilidad de que en un día la demanda sea superior a la potencia disponible?

Solución:

a) Cada central puede estar en dos estados, funcionando en cuyo caso produce su potencia nominal, o averiada que no produce nada. En la tabla se proporciona las ocho posibles combinaciones junto con sus probabilidades:

Tabla 3.4: Potencia producida

Nuclear	Térmica	Hidráulica	Potencia	Probabilidad
0	0	0	0	$0{,}10 \times 0{,}12 \times 0{,}05 = 0{,}0006$
0	0	200	200	$0{,}10 \times 0{,}12 \times 0{,}95 = 0{,}0114$
0	300	0	300	$0{,}10 \times 0{,}88 \times 0{,}05 = 0{,}0044$
500	0	0	500	$0{,}90 \times 0{,}12 \times 0{,}05 = 0{,}0054$
0	300	200	500	$0{,}10 \times 0{,}88 \times 0{,}95 = 0{,}0836$
500	0	200	700	$0{,}90 \times 0{,}12 \times 0{,}95 = 0{,}1026$
500	300	0	800	$0{,}90 \times 0{,}88 \times 0{,}05 = 0{,}0396$
500	300	200	1000	$0{,}90 \times 0{,}88 \times 0{,}95 = 0{,}7524$

La función de distribución de X la potencia disponible se deduce de la tabla y es igual a

$$F_X(x) = \begin{cases} 0 & \text{if } x < 0 \\ 0{,}0006 & \text{if } 0 \leq x < 200 \\ 0{,}0120 & \text{if } 200 \leq x < 300 \\ 0{,}0164 & \text{if } 300 \leq x < 500 \\ 0{,}1054 & \text{if } 500 \leq x < 700 \\ 0{,}2080 & \text{if } 700 \leq x < 800 \\ 0{,}2476 & \text{if } 800 \leq x < 1000 \\ 1{,}0000 & \text{if } x \geq 1000 \end{cases}$$

b) Sean U, V y W el número de centrales nucleares, térmicas e hidráulicas respectivamente que funcionan un día determinado. Las tres son variables aleatorias binomiales con $n = 30$ y p igual a 0.10, 0.12 y 0.05, respectivamente. Interesa conocer la distribución de probabilidad de Y,

$$Y = 500U + 300V + 200W,$$

que se aproxima a una normal de media

$$\begin{aligned} E[Y] &= 500E[U] + 300E[V] + 200E[W] \\ &= 500 \times 30 \times 0{,}9 + 300 \times 30 \times 0{,}88 + 200 \times 30 \times 0{,}90 \\ &= 27120 \end{aligned}$$

y varianza

$$\begin{aligned} Var[Y] &= 500^2 Var[U] + 300^2 Var[V] + 200^2 Var[W] \\ &= 500^2 \times 30 \times 0{,}9 \times 0{,}10 + 300^2 \times 30 \times 0{,}88 \times 0{,}12 + 200^2 \times 30 \times 0{,}90 \times 0{,}10 \\ &= 734851{,}2 \end{aligned}$$

por tanto

$$\begin{aligned} P(Y \leq 24000) &= P\left(\frac{Y - 27120}{857{,}2} \leq \frac{24000 - 27120}{857{,}2}\right) \\ &= P(Z \leq -3{,}639) = 0{,}0001368 \end{aligned}$$

c) Sea D la variable aleatoria "demanda", se pide calcular $P(W < D) = P(W - D < 0)$, la variable $S = W - D$ tiene distribución normal de media $27120 - 23000 = 4120$ y desviación típica $\sqrt{857{,}2^2 + 1000^2} = 1317{,}2$,

$$\begin{aligned} P(W < D) &= P(W - D < 0) \\ P(S \leq 0) &= P\left(\frac{S - 4120}{1317{,}2} \leq \frac{0 - 4120}{1317{,}2}\right) \\ &= P(Z \leq -3{,}13) = 0{,}000879 \end{aligned}$$

• • •

47. Una compañía desea aplicar un plan de muestreo para controlar la compra de lotes de 10000 unidades. La capacidad de inspección máxima que tienen es de 200 piezas. Determinar c, "*el número máximo de piezas defectuosas en la muestra que debe tener un lote aceptado*" si se desea que la probabilidad de rechazar un lote con el 10 % de piezas defectuosas sea igual 0.02. Un proveedor estima que sus lotes tienen un 5 % de piezas defectuosas, ¿qué probabilidad tiene de que un lote suyo sea aceptado?

Solución:

Sea X el número de defectuosas en la muestra. X sigue una distribución binomial con media $E[X] = 200 \times 0{,}1 = 20$ y varianza, $\mathrm{Var}[X] = 200 \times 0{,}1 \times 0{,}9 = 18$. Se debe obtener el valor de c que cumple $P(X \leq c) = 0{,}98$. Utilizando la aproximación normal,

$$P\left(\frac{X-20}{\sqrt{18}} \leq \frac{c-20}{\sqrt{18}})\right) = 0{,}98$$

En las tablas de la normal estndard se obtiene 2.05, que cumple $P(Z \leq 2{,}05) = 0{,}98$, que nos permite calcular c,

$$c = 20 + 2{,}05 \times \sqrt{18} \simeq 29$$

El plan es tomar una muestra de tamaño $n = 200$ y rechazar si el número de piezas defectuosas es mayor que 29.

La probabilidad de aceptar un lote con el 5 % de piezas defectuosas es:

$$\begin{aligned} P(X \leq 29{,}5) &= P\left(\frac{X - 200 \times 0{,}05}{\sqrt{200 \times 0{,}05 \times 0{,}95}} \leq \frac{29{,}5 - 200 \times 0{,}05}{\sqrt{200 \times 0{,}05 \times 0{,}95}}\right) \\ &= P(Z \leq 6{,}32) = 1. \end{aligned}$$

• • •

48. Se tienen dos variables aleatorias independientes U y V, ambas con distribución uniforme en [0,1]. A partir de ellas se definen las variables aleatorias Z y W:

$$\begin{cases} Z = \sqrt{-2\log U} \\ W = 2\pi V \end{cases}$$

   a) Calcular las probabilidades $P(Z < z_0)$ y $P(W < w_0)$. Utilizar estos resultados para deducir la expresión de la función de densidad de probabilidad conjunta $f_{Z,W}(z,w)$ para las variables aleatorias Z y W, así como el domino donde están definidas.

b) Una vez caracterizadas las variables aleatorias Z y W por su función de densidad de probabilidad conjunta $f_{Z,W}(z,w)$ se vuelve a definir 2 nuevas variables aleatorias X e Y mediante la siguiente transformación:$\left\{\begin{array}{l} X = Z \cdot \cos W \\ Y = Z \cdot senW \end{array}\right.$

Para estas nuevas variables se comprueba que la función de densidad de probabilidad conjunta es:

$$f_{X,Y}(x,y) = \frac{1}{2\pi}\exp\left(-\frac{1}{2}\left(x^2+y^2\right)\right), \text{ con } (x,y) \in \mathbb{R}^2.$$

Demostrar que las variables aleatorias X e Y son independientes. Calcular la media y la varianza de las variables y sus funciones de densidad marginal identificando el tipo de variables de las que se trata.

c) Con la información del apartado (a) y (b) obtener el valor numérico de las siguientes probabilidades:

- $P(-1 < X < 1, -1 < Y < 1)$
- $P(X^2 + Y^2 < 1)$
- $P(Y < X)$

d) Finalmente, calcular los coeficientes de la siguiente tansformación lineal

$$\begin{pmatrix} X' \\ Y' \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} c & 0 \\ d & e \end{pmatrix}\begin{pmatrix} X \\ Y \end{pmatrix}$$

con $c, e > 0$, para que las nuevas variables aleatorias $(X', Y')^T$ tengan una distribución de probabilidad con vector de medias y matriz de varianzas dadas por:

$$\mu = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \qquad S = \begin{pmatrix} 4 & 2 \\ 2 & 5 \end{pmatrix}.$$

Solución:

a)

$$\begin{aligned} P(Z \le z) &= P(\sqrt{-2\log U} \le z) \\ &= P(U \ge e^{-\frac{1}{2}z^2}) \\ &= 1 - e^{-\frac{1}{2}z^2}, \quad , z \ge 0 \end{aligned}$$

$$\begin{aligned} P(W \le w) &= P(2\pi V \le w) \\ &= P\left(V \le \frac{w}{2\pi}\right) \\ &= \frac{w}{2\pi}, \quad 0 \le w \le 2\pi \end{aligned}$$

Derivando las funciones anteriores se obtienen las funciones de densidad de Z y W,

$$f_Z(z) = ze^{-\frac{1}{2}z^2}, \quad z > 0$$

$$f_W(w) = \frac{1}{2\pi}, \quad 0 \leq w \leq 2\pi$$

y la función de densidad conjunta es

$$f_{Z,W}(z,w) = \frac{z}{2\pi}e^{-\frac{1}{2}z^2}, \quad z > 0,\ 0 \leq w \leq 2\pi.$$

b) Este problema presenta un método muy conocido para generar números aleatorios con distribución normal a partir de dos números aleatorios con distribución uniforme. La función de densidad conjunta se puede poner como el producto de las don funciones de densidad marginales

$$f_{X,Y}(x,y) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}x^2\right)\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}y^2\right)$$

$f_X(x) = \dfrac{1}{\sqrt{2\pi}}\exp\left(-\dfrac{1}{2}x^2\right), \quad f_Y(y) = \dfrac{1}{\sqrt{2\pi}}\exp\left(-\dfrac{1}{2}y^2\right).$

X e Y son variables aleatorias normales, independientes de media cero y varianza unidad.

c) Las probabilidades pedidas se obtienen de la siguiente manera:

$P(-1 < X < 1, -1 < Y < 1) = P(-1 < X < 1)P(-1 < Y < 1) = 0{,}6827^2 = 0{,}4661$. Se ha tenido en cuenta que X e Y son N(0,1).

2) $P(X^2 + Y^2 < 1) = P(Z^2 < 1) = P(Z < 1) = 1 - e^{-\frac{1}{2}} = 0{,}3935$

3) $P(Y < X) = P(Y - X < 0) = 0{,}5$ Pues $Y - X \sim N(0, \sqrt{2})$

d) Calculando el vector de media y la matriz de varianzas de las nuevas variables e igualando a los valores proporcionados se obtienen los siguientes resultados:

$$E\begin{pmatrix} X' \\ Y' \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} c & 0 \\ d & e \end{pmatrix} E\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

como $E[X] = E[Y] = 0$, $a = 1$ y $b = 3$.

Por otra parte,

$$\begin{aligned} Var\begin{pmatrix} X' \\ Y' \end{pmatrix} &= \begin{pmatrix} c & 0 \\ d & e \end{pmatrix} Var\begin{pmatrix} X \\ Y \end{pmatrix} \begin{pmatrix} c & d \\ 0 & e \end{pmatrix} \\ &= \begin{pmatrix} c & 0 \\ d & e \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} c & d \\ 0 & e \end{pmatrix} \\ &= \begin{pmatrix} c & 0 \\ d & e \end{pmatrix}\begin{pmatrix} c & d \\ 0 & e \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 2 & 5 \end{pmatrix} \end{aligned}$$

de donde se deduce que $c = 2$, $d = 1$ y $e = 2$.

• • •

49. Hay un juego muy popular en EEUU que se juega en carnavales en los que un jugador paga 1 dólar, elige un número del 1 al 6 y lanza tres dados. La banca le paga tantos dólares como veces aparece el número elegido. En un día de carnaval participan en el juego 500 personas. Obtén la distribución de probabilidad de las ganancias de la banca para ese día, indicando su media y desviación típica.

    Solución:

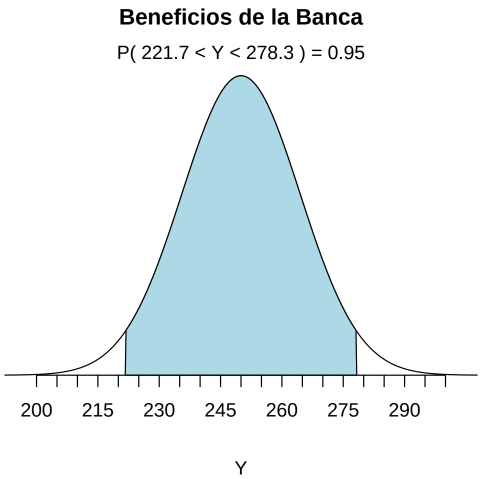


Figura 3.5: Ganancias en el juego de los dados

El número de aciertos X_k del jugador k es una variable binomial de parámetros $n = 3$ y $p = 1/6$. Los beneficios Y_k de la banca con el jugador k por tanto es $Y_k = 1 - X_k$, las ganancias con 500 jugadores es la variable aleatoria

$$Y = Y_1 + Y_2 + \cdots + Y_{500}$$

que según el teorema central del límite, tiene distribución normal de media

$$E[Y] = 500 \times (1 - 3 \times \frac{1}{6}) = 250$$

y varianza

$$Var[Y] = 500 \times 3 \times \frac{1}{6} \times \frac{5}{6} = 208{,}33$$

Los beneficios esperados son 250 dólares con una desviación típica de 14.5 dólares. Es un negocio con poco riesgo. En la figura (3.5) se representa la distribución de los beneficios, que con probabilidad del 95 % estarán entre 250 más menos 29 dólares.

• • •

50. Se lanzan 100 dados. Obtén la distribución de la suma de los números pares. Calcula la media y la varianza.

Solución:

Pensemos que los puntos que proporciona cada dado es el resultado del dado si el número es par y cero si el resultado del dado es impar.

Dado	Puntos	Probabilidad
1	0	1/6
2	2	1/6
3	0	1/6
4	4	1/6
5	0	1/6
6	6	1/6

En el dado k los puntos a sumar son $x = 0, 2, 4, 6$ tienen la siguiente distribución de probabilidad:

$$P(X_k = x) = \begin{cases} 1/2 & \text{if } x = 0\text{: Resultado impar} \\ 1/6 & \text{if } x = 2 \\ 1/6 & \text{if } x = 4 \\ 1/6 & \text{if } x = 6 \end{cases}$$

La media y la varianza de X_k es igual a:

$$E[X_k] = 0 \times \frac{1}{2} + 2 \times \frac{1}{6} + 4 \times \frac{1}{6} + 6 \times \frac{1}{6} = 2$$

$$\begin{aligned} Var[X_k] &= E[X_k^2] - E[X_k]^2 \\ &= 0^2 \times \frac{1}{2} + 2^2 \times \frac{1}{6} + 4^2 \times \frac{1}{6} + 6^2 \times \frac{1}{6} - 2^2 = \frac{16}{3} \end{aligned}$$

Si $X = X_1 + X_2 + \cdots + X_{100}$, entonces $E[X] = 200$ y $var[X] = 100\frac{16}{3} = 533{,}3$ (desviación típica $\simeq 23$). Por el teorema central del límite la distribución de la suma es normal.

DISTRIBUCIÓN NORMAL ESTÁNDAR $Z \backsim N(0,1) \to \Phi(z) = P(Z \leq z)$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.70884	0.71226	0.71566	0.71904	0.72240
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.75490
0.7	0.75804	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.86650	0.86864	0.87076	0.87286	0.87493	0.87698	0.87900	0.88100	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149	0.91309	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.92220	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.96080	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	0.97441	0.97500	0.97558	0.97615	0.97670
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.98030	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.98300	0.98341	0.98382	0.98422	0.98461	0.98500	0.98537	0.98574
2.2	0.98610	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.98840	0.98870	0.98899
2.3	0.98928	0.98956	0.98983	0.99010	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.99180	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.99430	0.99446	0.99461	0.99477	0.99492	0.99506	0.99520
2.6	0.99534	0.99547	0.99560	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.99720	0.99728	0.99736
2.8	0.99744	0.99752	0.99760	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.99900
3.1	0.99903	0.99906	0.99910	0.99913	0.99916	0.99918	0.99921	0.99924	0.99926	0.99929
3.2	0.99931	0.99934	0.99936	0.99938	0.99940	0.99942	0.99944	0.99946	0.99948	0.99950
3.3	0.99952	0.99953	0.99955	0.99957	0.99958	0.99960	0.99961	0.99962	0.99964	0.99965
3.4	0.99966	0.99968	0.99969	0.99970	0.99971	0.99972	0.99973	0.99974	0.99975	0.99976
3.5	0.99977	0.99978	0.99978	0.99979	0.99980	0.99981	0.99981	0.99982	0.99983	0.99983
3.6	0.99984	0.99985	0.99985	0.99986	0.99986	0.99987	0.99987	0.99988	0.99988	0.99989
3.7	0.99989	0.99990	0.99990	0.99990	0.99991	0.99991	0.99992	0.99992	0.99992	0.99992
3.8	0.99993	0.99993	0.99993	0.99994	0.99994	0.99994	0.99994	0.99995	0.99995	0.99995
3.9	0.99995	0.99995	0.99996	0.99996	0.99996	0.99996	0.99996	0.99996	0.99997	0.99997
4.0	0.99997	0.99997	0.99997	0.99997	0.99997	0.99997	0.99998	0.99998	0.99998	0.99998
4.1	0.99998	0.99998	0.99998	0.99998	0.99998	0.99998	0.99998	0.99998	0.99999	0.99999