TEMA1- DISTRIBUCIONES DE PROBABILIDAD FUNDAMENTALES.

Tras haber asimilado los temas anteriores sobre teoría de la probabilidad (contenidos de Estadística Empresarial I) se deben estudiar los modelos más "oportunos". Esta oportunidad se debe a las hipótesis de fenómenos reales de interés y a las distribuciones de estadísticos (que se generarán en el muestreo). Por su enorme variedad un estudio detallado de todos los modelos usados sería demasiado amplio y requeriría mucho más tiempo del disponible.

1.0.-REPASO DE MODELOS.

1.0.1.- DISTRIBUCIÓN BINOMIAL.

Se diferencia entre distribución binomial (1,p) y distribución binomial (n,p).

Distribución Binomial (1;p) o Distribución de Bernouilli.

Ejemplo. La distribución B(1;p) es la distribución de probabilidad discreta de los fenómenos aleatorios denominados dicotómicos como lanzar una moneda (interesa el número de caras obtenidas en ese lanzamiento) o como observar si un alumno supera una asignatura (interesa el número de aprobados para este caso único o individual).

Hipótesis. Fenómeno dicotómico (sólo dos sucesos posibles incompatibles y excluyentes) se puede concretar en dos sucesos, S y S^*; incompatibles, $S \cap S^* = \emptyset$; y cuya unión es el espacio muestral $S \cup S^* = E$ ya que ñabarcaò todas las posibilidades.

Construcción. En el modelo, un suceso (elegido arbitrariamente), por ejemplo S, se denomina suceso favorable o suceso éxito, se le asigna valor ñ1ò (uno) en la variable aletoria (al otro suceso que es el complementario complementario se le asigna valor ñ0ò) y probabilidad p (al complementario probabilidad ñqò con q=1-p):

$$\xi = \begin{cases} 1 & \text{si ocurre S} \\ 0 & \text{si no ocurre S} \end{cases} \quad \text{con} \quad \begin{cases} P(\xi = 1) = p \\ P(\xi = 0) = q = 1 - p \end{cases}.$$

Se dice, entonces, que una variante ξ tiene distribución de probabilidad binomial (1;p) si su función de cuantía es: $P(\xi = x) = p^x q^{1-x} \quad con \quad x = 0{,}1$. Esto significa que $P(\xi = 1) = p$ y $\mathrm{P}(\xi = 0) = \mathrm{q} = 1 - \mathrm{p}$.

La esperanza de una variante ξ con distribución binomial (1;p) es: *E(ξ)=* μ *= p.*

La varianza de una variante ξ con distribución binomial (1;p) es: *V(ξ) =* σ^2 *= pq.*

Este modelo uniparamétrico (un solo parámetro que es ñpò) se fija al evaluar el valor de ñpò y es muy útil dada la gran cantidad de fenómenos dicotómicos existentes y la posibilidad de obtener nuevos modelos en base a este tipo de fenómenos.

EJEMPLO.

*Se demuestra como ejemplo la esperanza y la varianza de la distribución binomial (1;p)**:***

$$\mathrm{E}(\xi) = \sum \mathrm{x_i} \cdot \mathrm{P}(\xi = \mathrm{x_i}) = 1 \cdot \mathrm{p} + 0 \cdot \mathrm{q} = \mathrm{p}$$

$$\mathrm{V}(\xi) = \mathrm{E}(\xi^2) - \mathrm{E}(\xi)^2 = \sum \mathrm{x_i}^2 \cdot \mathrm{P}(\xi = \mathrm{x_i}) - \mathrm{E}(\xi^2) = \mathrm{p} - \mathrm{p}^2 = \mathrm{p}(1 - \mathrm{p}) = \mathrm{pq}\,.$$

EJEMPLO.

Si en una urna hay 20 bolas blancas y 80 rojas, y se extrae una bola al azar, los dos únicos sucesos posibles son S_1 *=* ***ñ****Salir bola blanca****ò*** *y* S_2 *=* ***ñ****Salir bola roja****ò****.*

Se verifica que $P(S_1) = 0{,}2$ *y* $P(S_2) = 0{,}8$

La variable aleatoria $\xi = \begin{cases} 1 & \text{si sale blanca} \\ 0 & \text{si no sale blanca} \end{cases}$ *sigue una distribución binomial(1;0,2) o B(1;0,2). Su esperanza y su varianza se calculan fácilmente* $E(\xi) = p = 0{,}2 \quad y \quad V(\xi) = pq = 0{,}2 \times 0{,}8 = 0{,}16$.

<u>Distribución binomial (n;p).</u>

Ejemplo. La distribución B(n,p) es la distribución de probabilidad discreta que agrupa o suma fenómenos aleatorios dicotómicos independientes e iguales como lanzar diez veces una moneda (interesa el número de caras obtenidas en los diez lanzamientos) o como observar cuántos alumnos de un grupo superan una asignatura (interesa el número de aprobados en el grupo).

Hipótesis. Si se repite ñnò veces un mismo experimento aleatorio dicotómico, siendo las repeticiones independientes y con la misma probabilidad asignada al suceso éxito, entonces, la variable aleatoria ξ que representa el número de veces que sale el suceso denominado favorable o éxito sigue una distribución de probabilidad binomial(n;p).

Construcción. La distribución binomial (n;p) se obtiene como suma de ñnò variables aleatorias independientes con distribución de probabilidad binomial(1;p): $\xi=\xi_1+\xi_2+...+\xi_n$ tiene distribución de probabilidad binomial(n;p) si $\xi_1, \xi_2,..., \xi_n$ son n variables aleatorias independientes e iguales con distribución binomial(1;p).

Como consecuencia de lo anterior, una variable aleatoria ξ tiene distribución de probabilidad binomial (n;p) si su función de cuantía es:

$$P(\xi = i) = \binom{n}{i} \cdot p^i \cdot q^{n-i} \text{ para } i = 0, 1, ...,n.$$

Para evitar cálculos tediosos, los valores de las probabilidades están tabulados (hay tablas estadísticas con los valores de las probabilidades para diferentes combinaciones de parámetros) y se pueden obtener mediante sencillos habituales programas informáticos (como hojas de cálculo).

La esperanza de una variante ξ con distribución binomial (n;p) es: $E(\xi)= \mu =n\, p$.

La varianza de una variante ξ con distribución binomial (1;p) es: $V(\xi) = \sigma^2 = npq$.

EJEMPLO.

*Se demuestra como ejemplo la esperanza y la varianza de la distribución binomial (n;p)***:**

$$E(\xi)=\mu =E(\xi_1+\xi_2+...+ \xi_n)= E(\xi_1)+E(\xi_2)+...+E(\xi_n)= np$$

$$V(\xi) = \sigma^2 = V(\xi_1+\xi_2+...+ \xi_n)= V(\xi_1)+V(\xi_2)+...+V(\xi_n)= npq.$$

EJEMPLO.

Se analiza el fenómeno número de bolas blancas extraídas si se realizan 10 extracciones con reemplazamiento de una urna que contiene 20 bolas blancas y 80 rojas.

*Se repite, por tanto, el mismo experimento n = 10 veces. El suceso que llamamos favorable es S_1 = **ñ**Salir bola blanca**ò** con probabilidad p = 0,2 = P(S_1).*

*La variable aleatoria ξ = **ñ**Nº de bolas blancas de las 10 extracciones**ò** puede tomar con probabilidad diferente a cero sólo los valores i=0,1, 2,...,10 y tiene distribución de probabilidad binomial(10;0,2) o B(10;0,2), con función de cuantía:*

$$P(\xi = i) = \binom{10}{i} \cdot 0{,}2^{i} \cdot 0{,}8^{10-i}, \; i = 1,\, 2,\, 3,...,10.$$

La probabilidad, por ejemplo, de extraer exactamente 3 bolas blancas (o igualmente misma situación, exactamente 7 bolas negras) será:

$$P(\xi = 3) = \binom{10}{3} \cdot 0{,}2^{3} \cdot 0{,}8^{7} = \frac{10!}{3! \cdot 7!} \cdot 0{,}2^{3} \cdot 0{,}8^{7} = 0{,}2013.$$

Su esperanza y su varianza se calculan fácilmente $E(\xi) = n \cdot p = 10 \times 0{,}2 = 2$ *y* $V(\xi) = n \cdot p \cdot q = 10 \times 0{,}2 \times 0{,}8 = 1{,}6$.

Este modelo biparamétrico (dos parámetros ñnò y ñpò) se fija al evaluar el valor de los parámetros y es muy útil porque construye una variable ñcontadorò de sucesos éxito sobre un grupo de fenómenos dicotómicos repetidos independientemente.

La distribución binomial (n;p) tiene la propiedad aditiva o reproductiva en el parámetro p, esto es, la suma de variables aleatorias independientes con distribución binomial (n_i;p) también tiene una distribución binomial (mismo común parámetro p, suma de los diferente parámetros n_i).

1.0.2.- DISTRIBUCIÓN POISSON.

Ejemplo. La distribución Poisson (ə) o Ley de los Casos Raros es una distribución de probabilidad discreta que puede agrupar muchos (infinitos en teoría) fenómenos aleatorios dicotómicos independientes e iguales pero con probabilidad del suceso relevante o ñéxitoò muy pequeña, como observar cuántos alumnos de un grupo muy muy grande obtienen Matrícula de Honor con máxima calificación en una asignatura difícil (interesa el número de Matrículas de Honor en este numeroso grupo).

Hipótesis. Se puede determinar esta distribución de Poisson (también denominada de los casos raros) como límite de la distribución binomial (n;p) si se plantea (partiendo de la distribución binomial con sus hipótesis correspondientes) que el número de repeticiones del suceso dicotómico tiende a infinito (a nivel práctico u operativo más de 100), la probabilidad del suceso éxito tiende a cero (a nivel práctico u operativo menos de 0,05) y el producto de número de repeticiones por probabilidad del suceso éxito es un valor real (a nivel práctico u operativo menor de 20).

Construcción. Así planteada, la distribución Poisson de parámetro ə se obtiene como suma de n (con nŸ Đ) variables aleatorias independientes con distribución de probabilidad binomial (1;pŸ 0) y siendo n·p valor real: $\xi=\xi_1+\xi_2+...+\xi_n$ tiene distribución de probabilidad Poisson de parámetro ə= n·p si $\xi_1, \xi_2,..., \xi_n$ son nŸ Đ variables aleatorias independientes con distribución binomial(1;pŸ 0).

Como consecuencia de lo anterior, una variable aleatoria ξ tiene distribución de probabilidad Poisson de parámetro ə si su función de cuantía es:

$$P(\xi = i) = \frac{e^{-\lambda}\lambda^{i}}{i!}, \text{ para } i = 0, 1, ... ,Đ.$$

Aunque en teoría en la distribución de probabilidad de Poisson la variante puede tomar (con probabilidad no cero) cualquier valor natural más el cero (nunca valores negativos o no enteros), la forma de la función de cuantía indica que las probabilidades de que la variable tome valores cada vez mayores decrecen rápidamente. Esto confirma la tipología de fenómenos que se adaptan o son representados por este modelo: fenómenos con sucesos que aun siendo posibles son muy poco probables, el fenómeno se tiene que repetir muchas veces para que se presenten (sucesos raros como siniestros, defectos en producción en serie, partos múltiples, etc.).

Para evitar cálculos tediosos, los valores de las probabilidades están tabulados (hay tablas estadísticas con los valores de las probabilidades para diferentes valores del parámetro) y se pueden obtener mediante sencillos habituales programas informáticos (como hojas de cálculo).

La esperanza de una variante ξ con distribución Poisson (ə) es: *E(ξ)=* μ = ə.

La varianza de una variante ξ con distribución Poisson (ə) es: *V(ξ)* = σ^2 = ə.

A veces, esta coincidencia entre valor de esperanza y varianza puede ser un ñindicioò ante un problema para ñreconocerò un distribución de Poisson.

EJERCICIO.

El encargado de una cadena de montaje sospecha que la probabilidad de que un día salgan ***ñ****r****ò*** *(r=0,1,2,3,****é*** *) artículos defectuosos sigue una distribución con esperanza y varianza igual a cuatro, ya que la probabilidad de que un artículo sea defectuoso es 0ô004 y diariamente se producen 1.000 unidades. ¿Bajo qué hipótesis se justificaría la sospecha del encargado?. Probabilidad de que en un determinado día no haya artículos producidos defectuosos.*

Podría concluirse que se puede modelar con la ley de Poisson si se admite: que cada elemento o artículo producido respecto a tener defecto es un fenómeno dicotómico (con sólo dos sucesos incompatible y disjuntos); si se plantea la agrupación de todos los artículos producidos en el día y se cuentan los artículos defectuosos con una B(n,p) porque se supone suma, independencia e igual distribución de la variante dicotómica agregada; y, finalmente, si se admite que la producción diaria es suficientemente grande, que la probabilidad de que un artículo sea defectuoso es suficientemente pequeña y el producto de numero por probabilidad es estable se llega a la ley de Poisson.

Con ə*= n·p=1000·0,4=4 se tiene* $P(\xi = i) = \frac{e^{-4}4^{i}}{i!}$, para *i = 0, 1, ... ,*Ð *y*

$P(\xi = 0) = \frac{e^{-4}4^{0}}{0!}$ *=0,018315.*

EJERCICIO PROPUESTO.

Se contrata un seguro de vida para cada uno de los 20.000 deportistas federados de una región. Es razonable esperar que la probabilidad de que cualquiera *de estos deportistas fallezca es el 0,001% para un año. Determinar el número de indemnizaciones por muerte durante el año que previsiblemente tendrá que satisfacer la compañía de seguros y la probabilidad de que se produzcan dos o menos fallecimientos en el año.*

EJERCICIO PROPUESTO.

El responsable de unos multicines sabe que el 0'5% de los espectadores pasan sin entrada. Si se revisa al azar a 1000 espectadores, ¿cuál es la probabilidad de que más de tres hayan pasado sin entrada?. Si por fallos en la seguridad el 10% pasa sin entrada, tomando a 1000 espectadores, ¿cuál es la probabilidad de que menos de noventa hayan pasado sin entrada?.

Este modelo uniparamétrico (un parámetro λ) proporciona una variable ñcontadorò de casos raros: sucesos éxito muy poco probables o de muy baja posibilidad de acaecimiento sobre un grupo muy grande de fenómenos dicotómicos repetidos independientemente.

La distribución Poisson tiene la propiedad aditiva o reproductiva, esto es, la suma de variables aleatorias independientes con distribución Poisson de parámetro λ_i también tiene una distribución Poisson de parámetro λ (suma de los diferente parámetros $\lambda = \lambda_1 + \lambda_2 + ... + \lambda_k$).

EJERCICIO PROPUESTO.

Un determinado país está interesado en analizar el problema ecológico que supone el vertido de la carga de los petroleros. Tras estudiar los informes técnicos disponibles acepta las siguientes dos conclusiones. Primera, anualmente 1000 petroleros de doble casco realizan trayectos con carga y la probabilidad de vertido se estima en 0ô0001. Segunda, anualmente un número muy alto (pero desconocido) de petroleros monocasco realizan trayectos con carga y la probabilidad de siniestro no se ha calculado, si se sabe que al año hay, por término medio, 12 accidentes.

1.- Presentar y justificar un modelo que represente el número de vertidos anuales de buques de doble casco. Hallar el número esperado de vertidos anuales y la probabilidad de que no se produzca ningún vertido en un año.

2.- Presentar y justificar un modelo que represente el número de vertidos anuales de buques monocasco. Hallar la probabilidad de que no se produzca ningún vertido en un año y la probabilidad de que se produzcan menos de cinco vertidos.

3.- Presentar y justificar un modelo que represente el número de vertidos anuales de petroleros (tanto monocasco como de doble casco). Hallar el número esperado de vertidos anuales y la probabilidad de que se produzcan más de diez vertidos en un año.

1.0.3.- DISTRIBUCIÓN UNIFORME.

Ejemplo. La distribución Uniforme (a,b) o U(a,b) es una distribución de probabilidad continua que puede representar a fenómenos aleatorios sobre los que existe poca información y que presentan límites entre sus posibles valores, como ventas de componentes en una semana, sabiendo sólo que el número de ventas oscilará entre 10.000 y 20.000 unidades o como se emplea en la simulación por ordenador de los valores de una variante con una distribución determinada.

Hipótesis. La distribución uniforme surge a nivel más operativo cuando se considera una variable aleatoria con infinitos o ñmuchísimosò valores intermedios, que toma valores en un intervalo finito delimitado y de manera equiposible o equiprobable.

Construcción. Se dice que una variable aleatoria continua ξ tiene distribución de probabilidad uniforme en el intervalo [a,b] con -Ð<a<b<+Ð si su función de densidad viene dada por:

$$f(x) = \begin{cases} \dfrac{1}{b-a} & si \quad a \le x \le b \\ 0 & resto \end{cases}.$$

EJERCICIO PROPUESTO.

Demostrar que la función anterior es función de densidad.

Se citan a continuación algunas características relevantes.

1. Su función de distribución es: $F(x) = \begin{cases} 0 & si \quad x < a \\ \dfrac{x-a}{b-a} & si \quad a \le x \le b \\ 1 & si \quad x > b \end{cases}$.

2. Su esperanza es: $E(\xi) = \dfrac{a+b}{2}$.

3. Su varianza es: $V(\xi) = \dfrac{(b-a)^2}{12}$.

Evidentemente, al ser una distribución continua todas las probabilidades puntuales o P(ξ = x) son cero dado que las probabilidades ñse buscanò en los intervalos.

No presenta la U(a,b) la propiedad aditiva.

EJERCICIO PROPUESTO.

Demostrar estas características.

EJERCICIO.

Se considera un intervalo donde una variante se distribuye uniformemente con esperanza igual a ocho y varianza igual a nueve. Hallar función de densidad y la probabilidad de que la variante tome valores mayores a nueve.

Se necesita conocer la función de densidad, para ello se aprovecha la información sobre valores de la esperanza y de la varianza.

$$E(\xi) = \frac{a+b}{2} = 8 \ \ y \ \ V(\xi) = \frac{(b-a)^2}{12} = 9 \Rightarrow a=2{,}804 \ \ y \ \ b=13{,}196.$$

La función de densidad es $f(x) = \begin{cases} \dfrac{1}{10{,}392} & si \quad 2{,}804 \le x \le 13{,}196 \\ 0 & resto \end{cases}$

La probabilidad es $P(\xi > 9) = \int_9^{\infty} f(x)dx = \int_9^{13{,}196} \frac{1}{10{,}392} dx = 0{,}4038.$

1.1.- MODELO NORMAL.

La distribución normal es el modelo más importante por ser la que mejor se ajusta a muchos fenómenos reales, por ser buena aproximación (otros modelos bajo ciertas condiciones siguen un comportamiento normal y la agrupación o suma de muchos fenómenos no significativos en el conjunto e independientes también) y por ser base de distribuciones en el muestreo imprescindibles para la inferencia estadística. Es por esto que aparece la distribución normal en todos los campos de las ciencias empíricas.

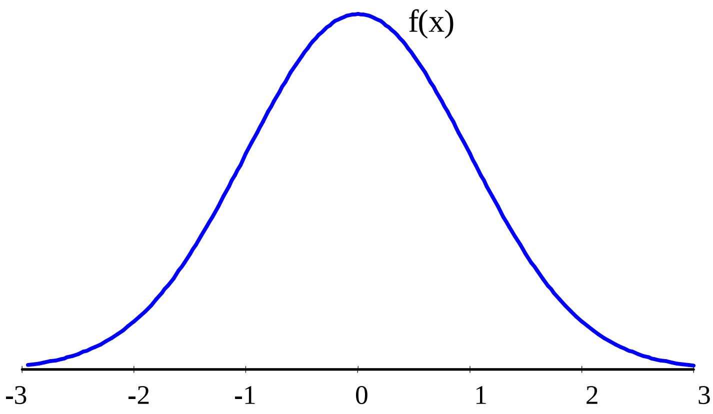
Para una más fácil comprensión de este modelo se empieza estudiando un caso particular (la normal reducida) para, después, continuar con la normal general.

Distribución Normal (0;1) o normal reducida o normal tipificada.

Hipótesis. Se dice que una variable aleatoria continua ξ tiene distribución de probabilidad normal con media 0 y desviación típica 1 o N(0,1) si su función de densidad viene dada por:

$$f(x)=\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{x^2}{2}} \quad con \quad -\infty<x<+\infty.$$

Características. La función de densidad se representa gráficamente dando lugar a lo que se conoce como ñcampana de Gaussò:



A partir de la función de densidad se demuestran interesantes propiedades que permiten entender mejor la distribución normal reducida:

1. La función de densidad es simétrica respecto al eje de ordenadas o recta x=0, esto es, f(-x)=f(x).

2. Media, mediana y moda se encuentran en el valor x=0.

3. Tiene una asíntota horizontal en y=0.

4. Alcanza un máximo absoluto en x=0 y es $f(x) = \dfrac{1}{\sqrt{2\pi}}$.

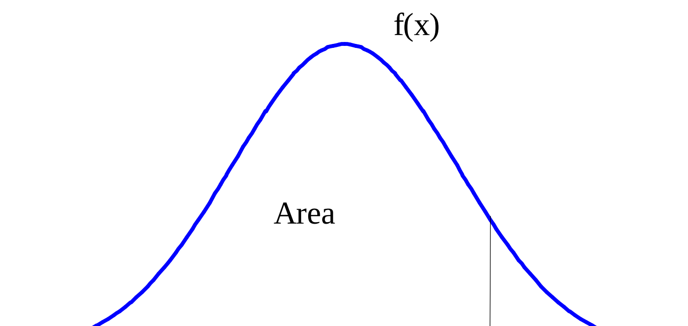
5. Es creciente para x<0 y decreciente para x>0.

6. Los puntos x=-1 y x=+1 son puntos de inflexión.

7. Mucha probabilidad en la zona central: en [-2,+2] probabilidad de 0,955 y en [-3,+3] probabilidad de 0,997.

Evidentemente, al ser una distribución continua todas las probabilidades puntuales o P(ξ = x) son cero dado que las probabilidades ñse buscanò en los intervalos. Las probabilidades en intervalos de la recta real pueden calcularse mediante la integral definida (área bajo la curva f), tal y como se plantea con cualquier variable aleatoria continua.

Por ejemplo, P(ξ ≤ x) es el área a la izquierda del punto x, bajo la curva f(x)



La función de densidad no tiene primitiva y para calcular probabilidades se deberían utilizar procedimientos aproximativos. Para facilitar la operatividad se dispone de tablas que permiten calcular probabilidades de variables aleatorias con distribución Normal(0;1). Hay tablas que presentan las probabilidades acumuladas en la cola izquierda respecto un punto ñxò o P(ξ<x) y tablas que presentan las probabilidades acumuladas en la cola derecha respecto un punto ñxò o P(ξ>x). Las tablas sólo dan información de una de las ñcolasò ya que los intervalos de cualquier otro tipo se pueden transformar para poder buscar en las tablas la probabilidad deseada (recurriendo a la simetría y al suceso complementario). Otra posibilidad es obtener las probabilidades mediante sencillos y habituales programas informáticos (como hojas de cálculo).

EJEMPLOS PARA PRACTICAR.

$$P(\xi \geq 0{,}47) = 0{,}3192$$

$$P(\xi \leq 0{,}54) = 1 - P(\xi \geq 0{,}54) = 1 - 0{,}2946 = 0{,}7054$$

$$P(\xi \leq -1{,}13) = P(\xi \geq 1{,}13) = 0{,}1292$$

$$P(\xi \geq -0{,}73) = 1 - P(\xi \leq -0{,}73) = 1 - P(\xi \geq 0{,}73) = 1 - 0{,}2327 = 0{,}7673$$

$$P(-1{,}96 \leq \xi \leq 2{,}12) = P(\xi \leq 2{,}12) - P(\xi \leq -1{,}96) = P(\xi \geq -1{,}96) - P(\xi \geq 2{,}12) =$$
$$= (1 - 0{,}0250) - 0{,}0170 = 0{,}958\,.$$

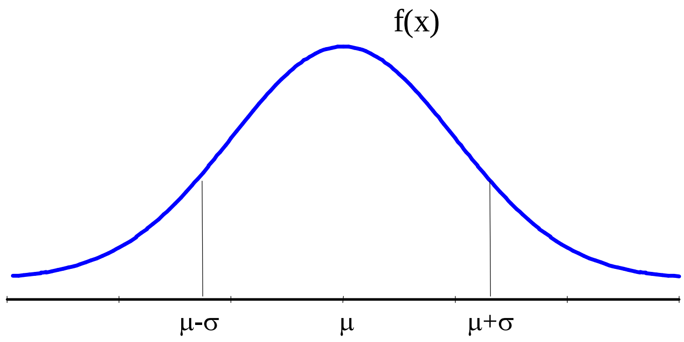
Distribución Normal(μ;σ) o Normal General.

Hipótesis. Se dice que una variable aleatoria continua ξ tiene distribución de probabilidad normal con media μ y desviación típica σ, si su función de densidad viene dada por:

$$f(x) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad con \quad \text{-Đ} < x < \text{+Đ}.$$

Dada una ξ*=N(0,1), se dice que una variable aleatoria continua ξ tiene distribución de probabilidad normal con media μ y desviación típica σ, ξ=N(μ,σ) si se obtiene como ξ=σ·ξ*+μ, es decir, N(μ,σ)=σ·N(0,1)+μ.

Características. La representación gráfica de esta función de densidad es:



A partir de la función de densidad se demuestran interesantes propiedades que permiten entender mejor la distribución normal general:

1. La función de densidad es simétrica respecto al eje de ordenadas o recta x=μ, esto es, f(μ-x)=f(μ+x).

2. Media, mediana y moda se encuentran en el valor x=μ.

3. Tiene una asíntota horizontal en y=0.

4. Alcanza un máximo absoluto en x=μ y es $f(x) = \frac{1}{\sigma \cdot \sqrt{2\pi}}$.

5. Es creciente para x<μ y decreciente para x>μ.

6. Los puntos x=μ-σ y x=μ+σ son puntos de inflexión.

7. Mucha probabilidad en la zona central: en [μ-2σ,μ+2σ] probabilidad de 0,955 y en [μ-3σ,μ+3σ] probabilidad de 0,997.

Evidentemente, al ser una distribución continua todas las probabilidades puntuales o P(ξ = x) son cero dado que las probabilidades ñse buscanò en los intervalos. Las probabilidades de intervalos corresponden, como se ha expuesto antes, a áreas bajo la curva función de densidad. Para calcular estas probabilidades nos podemos ayudar de las tablas de la distribución Normal(0;1) ya que cualquier variable aleatoria normal ξ se puede transformar en otra normal con media 0 y desviación típica 1, sin más que tipificar mediante la relación $P(\xi < x) = P(\sigma\xi * + \mu < x) = P(\xi* < \frac{x-\mu}{\sigma})$.

EJEMPLO.

La duración aleatoria ξ de un determinado componente eléctrico sigue una distribución normal de media 2.000 horas y desviación típica 100 horas. Se calcula la probabilidad de que un componente eléctrico elegido al azar dure entre 2.000 y 2.350 horas.

ξ es una variable aleatoria Normal(2.000;100). La variable aleatoria tipificada que tiene distribución Normal(0;1), a partir de la cual se pueden buscar las probabilidades en tablas de la normal es: $\xi^* = \dfrac{\xi - 2.000}{100}$.

La probabilidad que se busca resulta:

$$P(2.000 \le \xi \le 2.350) = P\left(\frac{2.000 - 2.000}{100} \le \frac{\xi - 2.000}{100} \le \frac{2.350 - 2.000}{100}\right) =$$

$$P(0 \le \xi^* \le 3{,}5) = P(\xi^* \ge 0) - P(\xi^* \ge 3{,}5) = 0{,}5 - 0{,}0002 = 0{,}4998.$$

La distribución Normal tiene la propiedad aditiva o reproductiva para combinaciones lineales, esto es, la combinación lineal de variables aleatorias independientes con distribución Normal también tiene una distribución Normal.

EJERCICIO PROPUESTO.

Obtener los parámetros resultantes de aplicar la propiedad reproductiva del modelo Normal.

<u>T.C.L. Teorema Central del Límite.</u>

Para terminar con los conceptos relacionados con la distribución Normal se tratará un aspecto de la convergencia de variables aleatorias. Un estudio profundo es de gran complejidad matemática y no se corresponde con los fines que se pretenden. Simplemente se destacará la importancia del comportamiento asintótico de la variable media aritmética: si la convergencia es en probabilidad se llega a una ley débil, si la convergencia es casi segura se llega a una ley fuerte y si la convergencia es en distribución hacia la normal se llega a los teoremas centrales del límite que serán el objeto de este apartado.

Como se dijo, un aspecto destacable del modelo Normal consiste en que es distribución límite de multitud de sucesiones de variables aleatorias, discretas y continuas. Esto se demuestra a través del teorema central del límite, que de manera intuitiva y operativa se puede describir de la siguiente forma:

ñSi tenemos una sucesión de variables aleatorias independientes (que se organiza como una suma de numerosas variables aleatorias), con la misma distribución y, con media y varianza finitas, la media aritmética de ellas (o la suma de las variantes) converge a una variable aleatoria con distribución normalò. Por ello, el teorema central de límite justifica que en determinadas condiciones sumas de variables aleatorias siguen una distribución aproximadamente normal.

Sean $\xi_1, \ldots \xi_n$, n variables aleatorias independientes igualmente distribuidas con media $E(\xi_j)$ y varianza $V(\xi_j)$ finita, si n ≥ 50 se puede escribir la variante suma ñSò como $P(S \leq x) = P(\xi_1 + \ldots + \xi_n \leq x) \approx P(N(\mu, \sigma) \leq x)$

siendo $E(S)=E(\xi_1+\xi_2+\ldots+\xi_n)= E(\xi_1)+E(\xi_2)+\ldots+E(\xi_n)= n\,E(\xi_j)=\mu$

$V(S)=V(\xi_1+\xi_2+\ldots+\xi_n)= V(\xi_1)+V(\xi_2)+\ldots+V(\xi_n)= n\,V(\xi_j)= \sigma^2$.

Si $\xi_j \equiv B(1, p) \Rightarrow S \equiv B(\sum_{j=1}^{n} 1, p) \equiv B(n, p) \cong N(np, \sqrt{npq})$,

si $\xi_j \equiv Poisson\ \lambda \Rightarrow S \equiv Poisson\ n\lambda \cong N(n\lambda, \sqrt{n\lambda})$,

si $\xi_j \equiv U(a,b) \Rightarrow S \equiv ???? \cong N\left(n\frac{a+b}{2}, \sqrt{\frac{(b-a)^2}{12}}\right)$.

EJERCICIO PROPUESTO.

Razonar que, siguiendo las ideas expuestas, para cualquier población con media finita μ y desviación típica σ, las distribuciones muestrales de la suma muestral y de la media muestral son aproximadamente normales si el tamaño n de la muestra es suficientemente grande.

EJERCICIO PROPUESTO.

Se lanza un dado perfecto. Hallar y justificar la elección del modelo utilizado para la probabilidad de obtener 2 resultados pares o menos al efectuar 10 lanzamientos y, para la probabilidad de obtener 6.000 pares o más al efectuar 10.000 lanzamientos.

1.2.- MODELO χ^2_n DE PEARSON.

Hipótesis. Si se tienen n variables aleatorias independientes, $\xi_1,...,\xi_n$, cada una de ellas con distribución Normal Reducida (media 0 y desviación típica 1), entonces la variable aleatoria $\xi = \xi_1{}^2 + ... + \xi_n{}^2$ recibe el nombre de χ^2 con n grados de libertad.

De otra forma, se dice que una variable aleatoria continua ξ tiene distribución de probabilidad χ_n^2 de Pearson si se puede escribir como $\xi = \xi_1{}^2 + ... + \xi_n{}^2$, donde, $\xi_1,...,\xi_n$, son variantes Normales independientes de media 0 y desviación típica 1.

También se dice que una variable aleatoria continua ξ tiene distribución de probabilidad χ_n^2 de Pearson si si su función de densidad es:

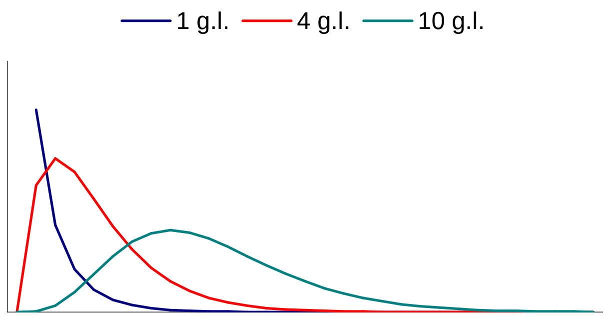
$$f(x) = \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}} \quad con \quad \text{-Ð}<x<+\text{Ð}.$$

Características. Es un modelos que ñse tieneò a partir de las distribuciones en el muestreo, no corresponde a un fenómeno real y por eso se dice que es, junto con otras distribuciones derivadas de la normal y resultado del muestreo, una distribución o modelo ñficticioò. Esta distribución es de carácter continuo, su función de densidad toma valores (dominio) en el intervalo $[0,\infty)$.

Su función de densidad será:

$$f(x) = \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}} \quad con \quad \text{-Ð}<x<+\text{Ð}.$$

La función de densidad depende de los grados de libertad, n. A modo de ejemplo se presentan en el siguiente gráfico la función de densidad para distintos grados de libertad:



El cálculo de probabilidades se realiza a través de tablas de la distribución χ^2. Algunas características son $E\left(\chi_n^2\right)=n,\quad V\left(\chi_n^2\right)=2n.$

La distribución χ^2 tiene la propiedad aditiva para suma de variantes, esto es, la suma de variables aleatorias independientes con distribución χ^2 también tiene una distribución χ^2: Dadas $\left\{X_i\right\}_{i=1}^{k} \to \chi_{n_i}^2$ independientes $\Rightarrow \sum_{i=1}^{k} X_i \to \chi_{n_1+n_2+\ldots+n_k}^2$.

En relación con la normal se puede exponer:

si $X \to \chi_n^2$, entonces $Y=\sqrt{2X} \xrightarrow{\approx} N(\sqrt{2n-1},1)$ si n es grande.

1.3.- MODELO t_n DE STUDENT.

Hipótesis. La distribución t_n de Student con n grados de libertad se puede construir a partir de n+1 variables aleatorias independientes con distribución N(0, σ), ξ_0, ξ_1,...ξ_n, de la siguiente forma:

$$\xi = \frac{\xi_0}{\sqrt{\frac{\xi_1^{\,2} + \ldots + \xi_n^{\,2}}{n}}} = \frac{N(0,1)}{\sqrt{\frac{\chi_n^2}{n}}} \ .$$

Se dice que una variable aleatoria continua ξ tiene distribución de probabilidad t_n de Student si su función de densidad es:

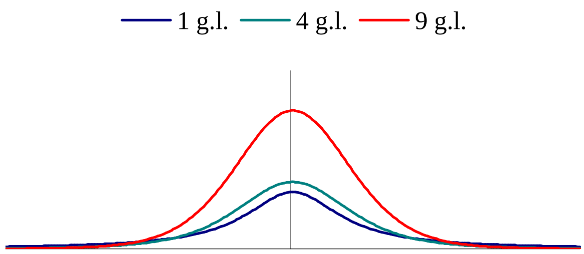
$$f(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\cdot\Gamma\left(\frac{n}{2}\right)}\cdot\left[1+\frac{x^2}{n}\right]^{-\frac{n+1}{2}} \quad con \quad \text{-Ð}<x<\text{+Ð}.$$

Características. Es un modelos que ñse tieneò a partir de las distribuciones en el muestreo, no corresponde a un fenómeno real y por eso se dice que es, junto con otras distribuciones derivadas de la normal y resultado del muestreo, una distribución o modelo ñficticioò. Esta distribución es de carácter continuo y toma valores en el intervalo $(-\infty,\infty)$.

Su función de densidad es:

$$f(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\cdot\Gamma\left(\frac{n}{2}\right)}\cdot\left[1+\frac{x^2}{n}\right]^{-\frac{n+1}{2}} \quad con \quad \text{-Ð}<x<\text{+Ð}.$$

Esta función es simétrica respecto al eje de ordenadas y su forma depende de los grados de libertad, n. A modo de ejemplo se comparan en el siguiente gráfico tres funciones diferentes:



La función de densidad resulta difícil de manejar, para calcular probabilidades se utilizan tablas. Otra posibilidad es obtener las probabilidades mediante sencillos y habituales programas informáticos (como hojas de cálculo).

Algunas características son: $E[T]=0$ para n>2 (si n=1 no existe media) y $Var[T]=\frac{n}{n-2}$ para n>3.

En relación con la normal se puede exponer que la distribución t de Student converge a una normal cuando los grados de libertad tienden a infinito (sirve para n>30).

1.4.- MODELO $F_{n,m}$ DE FISHER-SNEDECOR.

Hipótesis. Si se tienen n+m variables aleatorias independientes, cada una de ellas con distribución Normal Reducida (media 0 y desviación típica 1), entonces la variable aleatoria $F(m;n)=\dfrac{n\chi^2(m)}{m\chi^2(n)}$ recibe el nombre F de Fisher-Snedecor con m y n grados de libertad.

Se dice que una variable aleatoria continua ξ tiene distribución de probabilidad si su función de densidad es:

$$f(x)=\frac{m^{m/2}\cdot n^{n/2}\cdot\Gamma\left(\dfrac{m+n}{2}\right)}{\Gamma\left(\dfrac{m}{2}\right)\cdot\Gamma\left(\dfrac{n}{2}\right)}x^{\frac{m}{2}-1}(n+mx)^{-\frac{m+n}{2}} \quad con \quad -Đ<x<+Đ.$$

Características. Esta distribución es de carácter continuo, su función de densidad toma valores (dominio) en el intervalo $[0,\infty)$.

Su función de densidad es:

$$f(x)=\frac{m^{m/2}\cdot n^{n/2}\cdot\Gamma\left(\dfrac{m+n}{2}\right)}{\Gamma\left(\dfrac{m}{2}\right)\cdot\Gamma\left(\dfrac{n}{2}\right)}x^{\frac{m}{2}-1}(n+mx)^{-\frac{m+n}{2}} \quad con \quad -Đ<x<+Đ.$$

La función de densidad resulta difícil de manejar, para calcular probabilidades se utilizan tablas. Otra posibilidad es obtener las probabilidades mediante sencillos y habituales programas informáticos (como hojas de cálculo).