TEMAS DE MATEMÁTICAS
(Oposiciones de Secundaria)

TEMA 66

DISTRIBUCIONES DE PROBABILIDAD DE VARIABLE CONTINUA. CARACTERÍSTICAS Y TRATAMIENTO. LA DISTRIBUCIÓN NORMAL. APLICACIONES.

TEMA 66

DISTRIBUCIONES DE PROBABILIDAD DE VARIABLE CONTINUA. CARACTERÍSTICAS Y TRATAMIENTO. LA DISTRIBUCIÓN NORMAL. APLICACIONES.

1. VARIABLES ALEATORIAS CONTINUAS.

En el caso de variables discretas es posible asociar una probabilidad finita a cada punto (elemento) del conjunto muestral, aunque el número de puntos (elementos) que lo compongan sea infinito (numerable). La suma de todas las probabilidades será la unidad.

En el caso de una variable continua esto no es posible. Las probabilidades no sumarán uno, a menos que prácticamente a todos sus puntos (todos menos un conjunto numerable) les asignemos probabilidad cero. La solución a este problema que se plantea es considerar intervalos en lugar de puntos.

DEF Diremos que X es una Variable Aleatoria Continua (Unidimensional) si existe una función f(x) que verifica

1) $f(x) \geq 0 \quad \forall x \in \mathbb{R}$
2) $\int_{-\infty}^{+\infty} f(x)dx = 1$

y tal que para cualquier suceso A $\quad P(A) = P(X \in A) = \int_A f(x)dx$

La función f(x) recibe el nombre de función de densidad.

A partir de ahora consideraremos que las variables aleatorias continuas tienen función de densidad también continua, salvo, a lo sumo, en un número finito de puntos.

Definamos A como $A=\{x : a<x<b\}$. Entonces

$$P(A) = P(a < X < b) = \int_A f(x)dx = \int_a^b f(x)dx$$

Veamos algunas propiedades que verifican las variables aleatorias continuas.

1) Debido a que la integral anterior tiene el mismo valor independientemente de si los puntos a y/o b pertenecen al intervalo, tenemos que se verifica

$$P(a<X<b) = P(a<X\leq b) = P(a\leq X<b) = P(a\leq X\leq b)$$

2) A partir de la definición de probabilidad que hemos dado para las variables aleatorias continuas, se verifica que la probabilidad en un punto es cero

$$P(X = a) = \int_a^a f(x)dx = 0$$

3) Si A no es un intervalo sino la unión disjunta de un número finito de intervalos

$$A = A_1 \cup A_2 \cup ... \cup A_n \quad \text{con } A_i \cap A_j = \varnothing \ \ \forall i \neq j \quad \text{y} \ \ A_i = \{x : a_i < x < b_i\}$$

se verifica que

$$P(A) = P(A_1 \cup A_2 \cup \ldots \cup A_n) = \int_A f(x)dx = \int_{A_1} f(x)dx + \int_{A_2} f(x)dx + \ldots + \int_{A_n} f(x)dx =$$

$$= \int_{a_1}^{b_1} f(x)dx + \int_{a_2}^{b_2} f(x)dx + \ldots + \int_{a_n}^{b_n} f(x)dx$$

Esta propiedad establece que la probabilidad de que el valor de una variable aleatoria pertenezca a un conjunto A es el área comprendida entre f(x) y el eje x sobre el conjunto. Así se justifica la segunda condición de las funciones de densidad, donde en el caso de que el suceso A sea toda la recta real su probabilidad (según el axioma P2) debe ser la unidad.

Cualquier función f puede servir como función de densidad de una variable aleatoria continua siempre que verifique las dos condiciones dadas en la definición anterior.

Es claro que en un problema particular, se elegirá como función de densidad aquella que, para todo a y b (a<b)

$$P(a < X < b) = \int_a^b f(x)dx$$

represente la probabilidad de que el valor de la variable aleatoria X se halle comprendido entre a y b.

cualquier función positiva en un dominio elegido arbitrariamente puede considerarse como una función de densidad de una variable aleatoria, siempre que la función esté multiplicada por una constante que haga que su integral, en todo su dominio, sea 1. Así, por ejemplo

$$f(x) = \begin{cases} 0 & x \leq 2 \\ \dfrac{2x+3}{18} & 2 < x < 4 \\ 0 & x \geq 4 \end{cases}$$

es una función de densidad, ya que

$$\int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^{2} 0dx + \int_2^4 \frac{2x+3}{18}dx + \int_4^{+\infty} 0dx = 0 + 1 + 0 = 1$$

2. DISTRIBUCIONES MULTIVARIANTES.

DEF Diremos que X e Y son variables aleatorias continuas distribuidas conjuntamente si existe una función f tal que f(x,y)≥0 $\forall X,Y \in \mathbb{R}$ y tal que para cualquier suceso A

$$P(A) = P((X,Y) \in A) = \int_A f(x,y)dxdy$$

Vemos que es necesario que

1) f(x,y)≥0 $\forall X,Y \in \mathbb{R}$
2) $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dydx = 1$

La función f recibe el nombre de Función de Densidad de las variables aleatorias X e Y.

En el caso de k variables aleatorias continuas, la definición de distribución conjunta es análoga.

Ejemplo.

Definamos la función de distribución siguiente

$$f(x,y) = \begin{cases} \frac{1}{8}(6-x-y) & 0<x<2 \quad 2<y<4 \\ 0 & en \quad otro \quad caso \end{cases}$$

Triviamente podemos comprobar que verifica las dos condiciones para ser función de densidad.

Si X e Y son dos variables aleatorias que tienen esta densidad, la probabilidad de que estén en la región X<1 y Y<3 sería:

$$P(X<1, Y<3) = \int_{-\infty}^{1}\int_{-\infty}^{3} f(x,y)dydx = \int_0^1\int_2^3 \frac{6-x-y}{8}dydx = \frac{3}{8}$$

La probabilidad de que X+Y sea inferior a 3 es

$$P(X+Y<3) = \int_0^1\int_2^{3-x} \frac{6-x-y}{8}dydx = \frac{5}{24}$$

La probabilidad de que X<1 cuando se sabe que Y<3 es

$$P(X<1 \mid Y<3) = \frac{P(X<1, Y<3)}{P(Y<3)}$$

el numerador ya lo conocemos, veamos el valor del denominador.

$$P(Y<3)=\int_0^2\int_2^3 \frac{6-x-y}{8}\,dydx=\frac{5}{8}$$

Y entonces

$$P(X<1|Y<3)=\frac{P(X<1,Y<3)}{P(Y<3)}=\frac{\frac{3}{8}}{\frac{5}{8}}=\frac{3}{5}$$

3. DISTRIBUCIONES ACUMULATIVAS.

Puesto que en el caso de variables continuas las probabilidades vienen dadas por integrales, resulta a menudo conveniente considerar las integrales de las densidades con preferencia a las densidades mismas.

DEF Dada f(x) función de densidad de una variable aleatoria X, definimos la función F(x), llamada función de distribución Acumulativa, como

$$F(x)=P(X\le x)=\int_{-\infty}^{x} f(t)dt$$

PROP Si F(x) es una función de distribución acumulativa, verifica las siguientes propiedades:

1) F es una función no decreciente.
2) F(–∞) = 0
3) F(+∞) = 1
4) F es continua.

Dem.

La demostración es inmediata.

La función de densidad, si existe, puede hallarse a partir de la función de distribución acumulativa derivando F en los puntos donde tiene derivada. Es decir,

$$f(x)=\frac{dF(x)}{dx}$$

La probabilidad de que X esté en un cierto intervalo a<X≤b puede expresarse del siguiente modo con la función de Distribución Acumulativa

$$P(a<X\le b) = P(X\le b) - P(X\le a) = F(b) - F(a)$$

Para más de dos variables aleatorias, la función de distribución acumulativa se define de forma análoga al caso de una variable.

$$F(x,y)=P(X\le x,Y\le y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f(t_1,t_2)dt_2dt_1$$

siendo f(x,y) la función de densidad.

Dada la función de distribución acumulativa F(x,y), podemos hallar la función de densidad derivando F respecto de cada una de sus variables, en el supuesto, claro está, de que las derivadas existan.

$$f(x,y)=\frac{\partial}{\partial x}\frac{\partial}{\partial y}F(x,y)$$

La probabilidad de que (X,Y) esté en un rectángulo cualquiera

$$\begin{array}{c} a_1 < X < b_1 \\ a_2 < Y < b_2 \end{array}$$

podemos escribirla mediante la función acumulativa como

$$\begin{aligned} P(a_1 < X < b_1, a_2 < Y < b_2) &= P(X < b_1, Y < b_2) - P(X < a_1, Y < b_2) - \\ &- P(X < b_1, Y < a_2) + P(X < a_1, Y < a_2) = \\ &= F(b_1, b_2) - F(a_1, b_2) - F(b_1, a_2) + F(a_1, a_2) \end{aligned}$$

Estas distribuciones pueden complicarse mucho cuando se trata de más de dos variables aleatorias. De hecho, muchos problemas importantes de estadística permanecen sin resolver por la excesiva complicación de las integrales que hay que resolver.

4. DISTRIBUCIONES MARGINALES.

Cada distribución de más de una variable tiene asociadas varias distribuciones marginales. Sea f(x,y) la función de densidad correspondiente a dos variables aleatorias continuas. Puede ocurrir que sólo nos interese una de las variables, por ejemplo X. Buscaremos entonces una función que al integrarla sobre un intervalo $a<X<b$ nos dé la probabilidad de que X esté situada en ese intervalo. Cada uno de esos intervalos corresponde en el plano XY a una franja vertical. Por tanto

$$P(a < X < b) = \int_a^b \int_{-\infty}^{+\infty} f(x,y)\,dy\,dx$$

Cualquiera que sea la especificación de X, los límites de integración de Y son toda la recta real, de modo que podemos definir una función como

DEF Llamamos Función de Densidad Marginal con respecto a X a

$$f_1(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy$$

ya que, para cualquier par de valores a y b se verifica que

$$P(a < X < b) = \int_a^b f_1(x)\,dx$$

De forma análoga podemos definir

DEF Llamamos Función de Densidad Marginal con respecto a Y a

$$f_2(y) = \int_{-\infty}^{+\infty} f(x,y)dx$$

ya que, para cualquier par de valores a y b se verifica que

$$P(a < Y < b) = \int_a^b f_2(y)dy$$

La función de distribución marginal acumulativa se encuentra fácilmente a partir de la función de distribución acumulativa. Para dos variables, esta función acumulativa marginal es

$$F_1(x) = \int_{-\infty}^{x}\int_{-\infty}^{+\infty} f(x,y)dydx = \int_{-\infty}^{x} f_1(x)dx = F(x,+\infty)$$

Igualmente, de forma análoga obtenemos

$$F_2(y) = \int_{-\infty}^{y}\int_{-\infty}^{+\infty} f(x,y)dxdy = \int_{-\infty}^{y} f_2(y)dy = F(+\infty,y)$$

5. DISTRIBUCIONES CONDICIONALES.

Sea una función de densidad de dos variables aleatorias X e Y, f(x,y). Buscamos una función f(x|y) que nos dé la densidad de X cuando conozcamos Y. Es decir, una función tal que

$$P(a < X < b \,|\, Y) = \int_a^b f(x\,|\,y)dx$$

para valores cualesquiera a y b.

DEF Definimos la función de densidad condicional f(x|y) como

$$f(x\,|\,y) = \frac{f(x,y)}{f_2(y)} \qquad con \quad f_2(y) > 0$$

DEF Definimos la función de densidad condicional f(y|x) como

$$f(y\,|\,x) = \frac{f(x,y)}{f_1(x)} \qquad con \quad f_1(x) > 0$$

La función de densidad f(x|y) es una función de densidad de la variable X, siendo Y un parámetro que tendrá un valor numérico determinado para cada función de densidad condicional dada. Por tanto, la función f_2(c) debe considerarse como una constante. La función de densidad conjunta f(x,y) está representada por una superficie en el plano XY. Un plano perpendicular a XY, que corte a éste según la recta y=c, cortará a la superficie según la curva f(x,c). El área limitada por esta curva es

$$\int_{-\infty}^{+\infty} f(x,c)dx = f_2(c)$$

Por tanto, si dividimos f(x,c) por f_2(c), obtenemos una función de densidad que es precisamente f(x|c)

Las distribuciones condicionales se definen de manera análoga para las distribuciones con varias variables. Así, para cinco variables aleatorias que tengan como función de densidad f(x_1, x_2, x_3, x_4, x_5), la función de densidad condicional de X_1, X_3, X_5 para valores dados de X_2 y X_4 es

$$f(x_1, x_3, x_5 \mid x_2, x_4) = \frac{f(x_1, x_2, x_3, x_4, x_5)}{f_{24}(x_2, x_4)}$$

6. DISTRIBUCIÓN UNIFORME.

Comenzaremos el estudio de las distribuciones continuas con la más elemental de todas ellas.

DEF Diremos que una variable aleatoria X es uniforme en el intervalo (a,b) cuando su función de densidad de probabilidades es constante en él. Es decir

$$f(x) = \begin{cases} \dfrac{1}{b-a} & a < x < b \\ 0 & en \quad el \quad resto \end{cases}$$

Su función de distribución vendrá dada por

$$F(x) = \begin{cases} 0 & x \leq a \\ \dfrac{x-a}{b-a} & a < x < b \\ 1 & x \geq b \end{cases}$$

Los parámetros de la distribución son los extremos del intervalo a y b. Estos parámetros verifican las desigualdades

$$-\infty < a < b < +\infty$$

La media y la varianza de la distribución uniforme son

$$m = \frac{a+b}{2} \qquad s^2 = \frac{(b-a)^2}{12}$$

7. DISTRIBUCIÓN NORMAL.

Pieza clave en el estudio de la Estadística la constituye la distribución normal. Muchas distribuciones de frecuencias tienen aspectos parecidos, aunque describen fenómenos muy diferentes: pesos y alturas de una gran población, errores de medida, etc. Se apunta de esta manera el concepto de distribución normal. La dificultad de realizar un cálculo rápido de áreas bajo la curva normal, así como la imposibilidad de obtenerlas para todos los valores posibles de μ y σ, nos conduce a buscar transformaciones que obvien esta dificultad.

DEF Diremos que una variable aleatoria X que toma todos los valores reales tiene una distribución normal si su función de densidad de probabilidades es de la forma

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad -\infty < x < +\infty$$

Sus parámetros fundamentales son su media (μ) que puede tomar cualquier valor real, y su varianza (σ^2) que debe ser positiva.

Examinando la primera y segunda derivada de la función de densidad, podemos realizar un estudio de la misma, destacando las siguientes propiedades, que no vamos a demostrar por su sencillez.

PROP La moda de la distribución (máximo de la función de densidad) se alcanza en x=μ.

PROP La curva es simétrica respecto de un eje vertical que pasa por x=μ.

PROP La curva tiene sus puntos de inflexión en $x = \mu \pm \sigma$.

PROP Los valores de la media, la moda y la mediana coinciden.

PROP El eje de abscisas (y=0) es la asíntota horizontal de la curva (por ambos lados).

PROP El área comprendida entre la curva y el eje horizontal es 1.

Vamos a obtener ahora los valores de la media y la varianza de esta distribución.

PROP La Media y la Varianza de la distribución normal son μ y σ^2, respectivamente.

<u>Dem.</u>

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} x \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

Haciendo el cambio de variable

$$z = \frac{x-\mu}{\sigma}$$

$$dx = \sigma\, dz$$

obtenemos

$$E(X)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}(\mu+\sigma z)e^{-\frac{z^2}{2}}dz=\frac{\mu}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{z^2}{2}}dz+\frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}ze^{-\frac{z^2}{2}}dz=\mu$$

Analizando las dos integrales en las que ha quedado descompuesto la original, tenemos que la primera es μ multiplicada por una integral cuyo valor es 1 y la segunda es una integral impar en un dominio simétrico, siendo su resultado 0.

La varianza de la distribución normal viene dada por

$$E\left[(X-\mu)^2\right]=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}(x-\mu)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$$

y haciendo de nuevo el cambio de variable

$$z=\frac{x-\mu}{\sigma}$$

dx = σ dz

obtenemos

$$E\left[(X-\mu)^2\right]=\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}z^2e^{-\frac{z^2}{2}}dz$$

Aplicando ahora la integración por partes con

$$u=z \qquad du=dz$$
$$v=-e^{-\frac{z^2}{2}} \qquad dv=ze^{-\frac{z^2}{2}}dz$$

llegamos a

$$E\left[(X-\mu)^2\right]=\frac{\sigma^2}{\sqrt{2\pi}}\left(\left[-ze^{-\frac{z^2}{2}}\right]_{-\infty}^{+\infty}+\int_{-\infty}^{+\infty}e^{-\frac{z^2}{2}}dz\right)=\sigma^2(0+1)=\sigma^2$$

La curva de una distribución de probabilidad continua o función de densidad de probabilidad, se traza de manera que bajo la curva limitada por las abscisas $x=x_1$ y $x=x_2$, es igual a la probabilidad de que la variable aleatoria X tome valores comprendidos entre x_1 y x_2. Así tenemos

$$P(x_1<X<x_2)=\frac{1}{\sigma\sqrt{2\pi}}\int_{x_1}^{x_2}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$$

La transformación $z=\dfrac{x-\mu}{\sigma}$ nos proporciona valores de una distribución normal de media 0 y varianza 1, N(0,1). Veámoslo

$$P(x_1<X<x_2)=\frac{1}{\sigma\sqrt{2\pi}}\int_{x_1}^{x_2}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx=\frac{1}{\sqrt{2\pi}}\int_{z_1}^{z_2}e^{-\frac{z^2}{2}}dz=P(z_1<Z<z_2)$$

donde se observa que Z es una variable aleatoria normal de media 0 y varianza 1.

A la variable aleatoria que posee estas características se la denomina variable aleatoria normal estándar o tipificada y a su función de distribución, igualmente normal estándar.

Supongamos que X tiene una distribución N(0,1), entonces

$$P(a < X < b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{z^2}{2}} dz$$

Esta integral no puede evaluarse por métodos ordinarios. Por ello, debemos acudir a los métodos de integración numérica. La función de distribución de la normal estandarizada o tipificada es

$$\Phi(s) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{s} e^{-\frac{z^2}{2}} dz$$

Esta función se encuentra recogida en tablas. Para el cálculo de la probabilidad de que la variable aleatoria X esté en el intervalo (a,b), teniendo X una distribución normal N(0,1), haremos uso de las tablas.

$$P(a < X < b) = \Phi(b) - \Phi(a)$$

La importancia de esta tabulación radica en que si X es una variable aleatoria con distribución normal cualquiera $N(\mu,\sigma^2)$, la probabilidad anterior puede calcularse realizando la transformación

$$z = \frac{x - \mu}{\sigma}$$

con lo cual tipificamos la variable aleatoria X y a continuación utilizamos la tabulación mencionada. Es decir

$$P(a < X < b) = P\left(\frac{a-\mu}{\sigma} < Z < \frac{b-\mu}{\sigma}\right) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

Una conclusión de la definición de Φ es que

$$\Phi(-x) = 1 - \Phi(x)$$

Esta relación es muy útil, puesto que en la mayoría de las tablas la función Φ aparece tabulada sólo para valores positivos de la variable.

Finalmente, vamos a calcular la probabilidad

$$P(\mu - k\sigma < X < \mu + k\sigma)$$

siendo X una variable aleatoria con distribución normal cualquiera $N(\mu,\sigma^2)$

$$P(m - k s < X < m + k s) = P\left(-k < \frac{X - m}{s} < k \right) = \Phi(k) - \Phi(-k) = 2\Phi(k) - 1$$

Expresión en la que podemos observar que la probabilidad anterior es independiente de μ y σ, y que ésta sólo depende de k.

7.1. Aproximación de las Distribuciones Binomial y Poisson a la Normal.

Las probabilidades asociadas a experimentos binomiales, cuando n es pequeña, se obtienen con facilidad mediante la fórmula B(x;n,p) (ya vista en el tema anterior). Para valores grandes de n, por lo general las probabilidades binomiales se calculan con procedimientos de aproximación. Si X es una variable aleatoria binomial de parámetros n y p, entonces, si n es grande y ni p ni q son próximos a cero, podemos considerar que X sigue una distribución normal N(np, npq). Esta propiedad la conocemos como

TEOREMA. Teorema de De Moivre.

Si X tiene una distribución binomial de parámetros n y p, entonces la variable aleatoria tipificada

$$Z = \frac{X - np}{\sqrt{npq}}$$

tiende a una distribución normal estándar N(0,1) cuando $n \to \infty$.

OBS En general, podemos considerar como buena aproximación cuando $n>25$ y $npq>5$.

Debemos resaltar que estamos aproximando una distribución discreta por una continua, por tanto, resulta necesario establecer una corrección por continuidad. Se ha encontrado que la siguiente corrección para continuidad mejora la proximidad anterior.

$$\begin{aligned}
P(X = k) &\equiv P\left(k - \frac{1}{2} \le X \le k + \frac{1}{2} \right) \\
P(a < X < b) &\equiv P\left(a + \frac{1}{2} \le X \le b - \frac{1}{2} \right) \\
P(a < X \le b) &\equiv P\left(a + \frac{1}{2} \le X \le b + \frac{1}{2} \right) \\
P(a \le X < b) &\equiv P\left(a - \frac{1}{2} \le X \le b - \frac{1}{2} \right) \\
P(a \le X \le b) &\equiv P\left(a - \frac{1}{2} \le X \le b + \frac{1}{2} \right)
\end{aligned}$$

donde k, a y b son números enteros.

De la misma manera que se ha establecido una aproximación de la distribución binomial a la normal, podemos establecer una aproximación de la distribución de Poisson a la normal.

Sea X una variable aleatoria de Poisson P(λ) con λ grande, en la práctica $\lambda>25$. En estas condiciones, la variable tipificada

$$Z = \frac{X - \lambda}{\sqrt{\lambda}}$$

tiene aproximadamente una distribución normal estándar N(0,1) y, por tanto, podemos considerar que X sigue aproximadamente una distribución N(λ,λ). Aquí también será conveniente tener en cuenta la corrección por continuidad similar a la efectuada para la distribución binomial.

8. OTRAS DISTRIBUCIONES.

Como la distribución Normal no resuelve todos los problemas, tenemos que recurrir a otras distribuciones para resolver determinados problemas en los campos de la ingeniería y de la ciencia, como son la distribución Gamma, Exponencial, Wiebull y χ^2.

8.1. Distribución Gamma.

DEF Definimos la función Gamma como

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x}\, dx \qquad \alpha > 0$$

Integrando por partes con

$$u = x^{\alpha-1}$$
$$dv = e^{-x}\, dx$$

se obtiene la fórmula de recurrencia

$$\Gamma(\alpha) = (\alpha-1)\cdot\Gamma(\alpha-1)$$

En particular, $\Gamma(1)=1$ y si $\alpha=n$, siendo n un número positivo, la recurrencia anterior conduce a que

$$\Gamma(n) = (n-1)!$$

Algunas expresiones útiles para la función gamma son

PROP La Función Gamma verifica

1) Fórmula de los complementos: $\Gamma(\alpha)\Gamma(1-\alpha) = \dfrac{\pi}{\operatorname{sen} \alpha\pi}$ con $0<\alpha<1$

2) $\Gamma\left(\dfrac{1}{2}\right) = \sqrt{\pi}$

DEF Diremos que la variable aleatoria X tiene una Distribución Gamma, con parámetros α y β, si su función de densidad viene dada por

$$f(x;\alpha,\beta)=\begin{cases}\dfrac{x^{\alpha-1}e^{-\frac{x}{\beta}}}{\beta^{\alpha}\Gamma(\alpha)} & x>0,\alpha>0,\beta>0\\ 0 & en \quad el \quad resto\end{cases}$$

donde α se conoce como parámetro de forma y β como parámetro de escala.

Cuando α es entero positivo, la distribución Gamma también se conoce con el nombre de Distribución de Erlang, en honor al científico danés que la utilizó por primera vez en el año 1900.

PROP La media y la varianza de la distribución gamma son

$$E(X)=\alpha\beta \qquad Var(X)=\alpha\beta^2$$

Dem.

En efecto, el r-ésimo momento en torno al cero vale

$$m_1'=\beta\frac{\Gamma(\alpha+1)}{\Gamma(\alpha)}=\beta\frac{\alpha\Gamma(\alpha)}{\Gamma(\alpha)}=\alpha\beta=E(X)$$

Y para la varianza tenemos

$$m_2'=\beta^2\frac{\Gamma(\alpha+2)}{\Gamma(\alpha)}=\beta^2\frac{(\alpha+1)\alpha\Gamma(\alpha)}{\Gamma(\alpha)}=(\alpha+1)\alpha\beta$$

$$Var(X)=E(X^2)-\left[E(X)\right]^2=m_2'-m_1'^2=\alpha(\alpha+1)\beta-\alpha\beta=\alpha\beta^2$$

8.2. Distribución Exponencial.

Si particularizamos la distribución Gamma para α=1 obtenemos:

$$f(x;\beta)=\begin{cases}\dfrac{e^{-\frac{x}{\beta}}}{\beta} & x>0,\beta>0\\ 0 & en \quad el \quad resto\end{cases}$$

que es la función de densidad de una variable aleatoria X con distribución exponencial.

Esta función exponencial se caracteriza por el parámetro β, que representa el lapso promedio de tiempo que transcurre entre dos sucesos independientes de Poisson. Es decir, la distribución exponencial puede modelar el tiempo transcurrido entre dos sucesos consecutivos de Poisson (llegadas o prestaciones de servicios).

Su función de distribución es

$$F(x;\beta)=P(X\le x)=1-e^{-\frac{x}{\beta}}$$

PROP La media y la varianza de la función exponencial son

$$E(X)=\beta \qquad Var(X)=\beta^2$$

Dem.

Basta tomar α=1 en la media y varianza de Gamma para obtener los resultados.

8.3. Distribución de Weibull.

La distribución de Weibull fue establecida por el físico suizo del que lleva su nombre y al igual que la distribución exponencial, su función fundamental es en el estudio de la fiabilidad.

DEF Diremos que una variable aleatoria X tiene una distribución de Weibull si su función de densidad de probabilidad es de la forma

$$f(x;\alpha,\beta)=\begin{cases}\dfrac{\alpha x^{\alpha-1}e^{-\left(\frac{x}{\beta}\right)^{\alpha}}}{\beta^{\alpha}} & x>0,\alpha>0,\beta>0\\ 0 & en \quad el \quad resto\end{cases}$$

La distribución de Weibull es una familia de distribuciones que dependen de dos parámetros, el de forma, α, y el de escala, β. Podemos introducir un parámetro adicional reemplazando la variable aleatoria de Weibull X, por X–a, donde a es un parámetro de localización, que representa un valor umbral de tiempo de garantía.

PROP La media y la varianza de la distribución Weibull son

$$E(X)=\beta\Gamma\left(1+\frac{1}{\alpha}\right) \qquad Var(X)=\beta^2\left[\Gamma\left(1+\frac{2}{\alpha}\right)-\Gamma^2\left(1+\frac{1}{\alpha}\right)\right]$$

Dem.

El r-ésimo momento en torno al cero de una variable aleatoria que sigue una distribución de Weibull es

$$\mu_r'=\beta^r\Gamma\left(1+\frac{r}{\alpha}\right)$$

con lo cual

$$\mu_1'=\beta\Gamma\left(1+\frac{1}{\alpha}\right)=E(X)$$

y para la varianza tenemos

$$Var(X) = E(X^2) - [E(X)]^2 = m_2' - m_1'^2 = b^2\Gamma\left(1+\frac{2}{a}\right) - \left[b\Gamma\left(1+\frac{1}{a}\right)\right]^2 =$$

$$= b^2\left[\Gamma\left(1+\frac{2}{a}\right) - \Gamma^2\left(1+\frac{1}{a}\right)\right]$$

8.4. Distribución c^2 de Pearson.

Sean $X_1, X_2, ..., X_n$ n variables aleatorias N(0,1) independientes entre sí. La variable $Y = X_1^2 + X_2^2 + \ldots + X_n^2$ recibe el nombre de χ^2 de Pearson con n grados de libertad y se denota por c_n^2.

Por estar formada la variable aleatoria Y por suma de cuadrados resulta positiva, además su función de densidad de probabilidad es

$$f(x;n) = \begin{cases} \dfrac{x^{\frac{n}{2}-1} e^{-\frac{x}{2}}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} & x > 0 \\ 0 & en \quad el \quad resto \end{cases}$$

Una propiedad importante de la c_n^2 es que la suma de varias de ellas que sean independientes nos da como resultado otra cuyo número de grados de libertad es la suma de los grados de libertad de éstas.

PROP La media y la varianza de una distribución c_n^2 es

E(X) = n Var(X) = 2n

Dem.

El r-ésimo momento en torno al cero es

$$m_1' = b\frac{\Gamma\left(\frac{n}{2}+1\right)}{\Gamma\left(\frac{n}{2}\right)} = \frac{2n}{2}\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} = n = E(X)$$

y para la varianza

$$m_2' = 4\frac{\Gamma\left(\frac{n}{2}+2\right)}{\Gamma\left(\frac{n}{2}\right)} = 4\left(\frac{n}{2}+1\right)\frac{n}{2}\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} = n^2 + 2n$$

$$Var(X) = E(X^2) - [E(X)]^2 = m_2' - m_1'^2 = n^2 + 2n - n^2 = 2n$$

Integrando por partes, desde 2λ hasta +∞, la función de densidad de una distribución χ^2 con n=2(r+1) grados de libertad, se obtiene la siguiente relación entre las probabilidades de una χ^2 con un número par de grados de libertad y las de una distribución de Poisson

$$P\left(\chi^2_{2(c+1)} > 2\lambda\right) = \int_{2\lambda}^{+\infty} \frac{x^c e^{-\frac{x}{2}}}{2^{c+1} c!} dx = \sum_{r=0}^{c} \frac{\lambda^r e^{-\lambda}}{r!} = P\left(\Pi(\lambda) \le c\right)$$

donde se ha tenido en cuenta que para c entero es Γ(c+1)=c!. En ciertas ocasiones suele utilizarse esta relación para calcular probabilidades de Poisson

BIBLIOGRAFÍA RECOMENDADA.

Estadística Teórica. Aut. J.M.Doblado y M.C. Nieto. Edit. UNED

Introducción a la Estadística Teórica. Aut.: G Arnáiz. Edit.: Lex Nova

Estadística Teórica y Aplicada. Aut.: A. Nortes. Edit.: S. Rodríguez.

Introducción a la Probabilidad y la Medida (I). Aut.: P Zoroa y N. Zoroa. Edit.: Maior DM.

Introducción a la Teoría de la Estadística. Aut.: Mood, Graybill. Edit.: Aguilar.