Menú principal

Principales distribuciones de una variable aleatoria continua.

Esquema

1. Distribución Uniforme
2. Distribución Normal
3. Distribución Chi-cuadrado
4. Distribución t-Student
5. Distribución F de Fisher.
6. Distribución de Pareto.
7. Distribución Log-Normal
8. Distribución Beta
9. Distribución Gamma
10. Distribución de Erlang
11. Distribución de Weibull
12. Distribución exponencial negativa

Como se sabe la ley de probabilidad de una distribución continua X está bien definida si se conoce su función de densidad f(x), bien si se conoce su función de distribución F(x), teniéndose que

$$F(x) = P(X \le x) \qquad P(a < X \le b) = \int_a^b f(x)dx \qquad \int_{-\infty}^{+\infty} f(x)dx = 1$$

estando f(x) y F(x) relacionados por la expresión

$$F(x) = \int_{-\infty}^{x} f(x)dx \quad o \quad f(x) = \frac{dF(x)}{dx}$$

Distribución Uniforme.

Se dice que la variable aleatoria X sigue una distribución uniforma en el intervalo [a, b] cuando su función de densidad es

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

En esta distribución hay que observar que:

a) Depende de los parámetros a y b

b) Su función de distribución es

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

c) Su media y varianza son

$$\mu = \frac{a+b}{2}, \quad \sigma^2 = \frac{(b-a)^2}{12}$$

[Volver al principio]

Distribución Normal o de Laplace-Gauss

La variable continua X se dice que sigue una distribución normal de media μ y desviación típica σ si:

a) Puede tomar cualquier valor en el intervalo (-∞; +∞)

b) Su función de densidad es:

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

la gráfica de esta función tiene forma de campana

En la distribución normal hay que observar:

c) Depende de dos parámetros μ y σ

d) Su función de distribución es igual a

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \quad si -\infty < x < \infty$$

e) Su media y varianza son μ y σ^2

f) La función de densidad presenta un máximo en $x=\mu$, dos puntos de inflexión en $x=\mu-\sigma$ y $x=\mu+\sigma$ y tiene el eje OX como asíntota.

g) La curva f(x) es simétrica respecto de $x=\mu$ y por tanto su media, mediana y moda coinciden. Esta propiedad es muy utilizada en el cálculo de áreas mediante tablas.

h) Si una variable X_1 es $N(\mu_1, \sigma_1)$ y otra variable X_2 es $N(\mu_2, \sigma_2)$ independientes entre sí, entonces la nueva variable aleatoria $X=X_1 \pm X_2$ sigue también su distribución normal de media $\mu_1 \pm \mu_2$ y desviación típica $\sigma = \sqrt{\sigma_1^2 + \sigma_2^2}$, es decir, la variable X sigue una

$$N(\mu_1 \pm \mu_2, \sqrt{\sigma_1^2 + \sigma_2^2})$$

Esta propiedad se generaliza a n variables.

i) Si X es una $N(\mu, \sigma)$ la variable

$$Z = \frac{X-\mu}{\sigma} \, sigue\, una \quad N(0,1)$$

su función de densidad será

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

A la variable Z se le llama variable tipificada de X y a la curva de su función de densidad curva normal tipificada

En la distribución normal tipificada hay que observar que

j) No depende de ningún parámetro

k) Su media y varianza son 0 y 1.

l) La curva es simétrica respecto del eje z=0, tiene puntos de inflexión en z=-1 y z=1

Si X es una variable binomial de parámetros n y p, con n muy grande y con p y q no muy próximos a 0. Podemos considerar que la variable X sigue una distribución normal

$$N(np; \sqrt{npq})$$

y por tanto la variable

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

sigue una normal N(0, 1) para que esta transformación sea correcta, hay que tener en cuenta que X es discreta y Z continua, por lo que P(X=a)= al considerar que X es continua= $P(a - 0.5 \leq x \leq a + 0.5) y\ P(a < x < b) = P(a + 0.5 \leq x \leq b - 0.5)$ y si tenemos en cuenta $P(a \leq x \leq b) = P(a - 0.5 \leq x \leq b + 0.5)$

En general la transformación es aceptable cuando (p<0.5 y np>5) ó (q<0.5 y nq>5)

La distribución de la variable Z se encuentra tabulada, en las tablas aparecen el área bajo la curva de la normal a la derecha de un punto Z_α. Al punto Z_α se le llama punto crítico y al área bajo la curva a la derecha del punto: α se le llama nivel de significación. Solo se encuentran valores de Z_α mayores o iguales que 0 ó áreas α menores o iguales que 0.5. Los demás valores se obtiene por simetría.

La gran utilidad de la variable tipificada Z es que nos permita calcular áreas de cualquier variable con distribución normal.

Si X es N(μ, σ) entonces

$$P(a \leq x \leq b) = P(\frac{a - \mu}{\sigma} \leq \frac{X - \mu}{\sigma} \leq \frac{b - \mu}{\sigma}) = P(z' \leq Z \leq z)$$

Este proceso se denomina tipificación.

[Volver al principio]

Distribución $\aleph^2$de Pearson.

Sean $X_1, X_2, \ldots, X_n$, n variables aleatorias N(0, 1) independientes entre sí. La variable

$$\chi_n^2 = X_1^2 + X_2^2 + \ldots + X_n^2$$

recibe el nombre de $\aleph^2$ de Pearson con n grados de libertad. Su función de densidad es

$$f_{\chi^2}(x) = \begin{cases} \frac{1}{2^{\frac{n}{2}}\Gamma(n/2)} e^{-x/2} . x^{(x/2)-1} & 0 < x < \infty \\ 0 & x \leq 0 \end{cases} \quad \text{la función } \Gamma(p) = \int_0^\infty e^{-x} . x^{p-1} dx \quad p > 0$$

En esta distribución hay que observar:

a) Sólo puede tomar valores positivos por tratarse de una suma de cuadrados de n variables.

b) Depende sólo del parámetro n, por lo que no hay, como ocurría en la normal, una curva tipificada única

c) Su media y varianza son $\mu=n$, $\sigma^2=2n$

d) Si χ_{n1}^2 y χ_{n2}^2 son dos distribuciones con n_1 y n_2 grados de libertad, independientes entre sí, entonces la suma de las dos es también una χ^2 con n_1+n_2 grados de libertad:

$$\chi_{n_1+n_2}^2 = \chi_{n_1}^2 + \chi_{n_2}^2$$

esta propiedad se puede generalizar a n variables.

e) Al aumentar el número de grados de libertad esta distribución se aproxima asintóticamente a la normal. Para n>30 la variable

$$\sqrt{2\chi_n^2} \xrightarrow{(n\to\infty)} N(\sqrt{2n-1},1)$$

f) En el muestreo, si tomamos muestras de media $\bar{x}$ y desviación típica s de una población normal $N(\mu, \sigma)$, la variable

$$\chi_{n-1}^2 = (n-1)\frac{s^2}{\sigma^2}$$

con s^2=cuasivarianza, es una χ^2 de Pearson con n-1 grados de libertad. Esta última propiedad es muy utilizada en la estimación y en el contraste de hipótesis sobre la varianza poblacional σ^2.

g) Se encuentra la tabla de la χ^2 para valores de n<=30.

h) Para el uso de la tabla consideramos las áreas bajo la curva a la derecha de un punto $\chi^2_{\alpha;n}$ crítico. A la derecha de este punto, el área bajo la curva es igual a α (nivel de significación) en una χ^2 de Pearson con n grados de libertad. Se tiene:

$$P(\chi_n^2 \geq \chi_{\alpha;n}^2) = \alpha$$

y para áreas a la izquierda

$$P(\chi_n^2 \leq \chi_{\alpha;n}^2) = 1-\alpha$$

[Volver al principio]

Distribución t de Student.

Sean $X_1, X_2, \ldots, X_n, X$; n +1 variables aleatorias N(0, σ) independientes entre sí. La variable t_n:

$$t_n = \frac{X}{\sqrt{\frac{1}{n}\sum_{i=1}^{n} X_i^2}}$$

se denomina t de Student con n grados de libertad. Dividiendo en el segundo miembro numerador y denominador por σ, tenemos:

$$t_n = \frac{X/\sigma}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i^2/\sigma^2)}} = \frac{Z}{\sqrt{\frac{1}{n}\chi_n^2}}$$

con Z es N(0, 1) y χ_n^2 es una χ^2de Pearson con n grados de libertad.

La función de densidad de esta t de Stundent con n grados de libertad es:

$$f_{t_n}(x) = \frac{1}{\sqrt{n}\beta(1/2, n/2)}.(1+\frac{x^2}{n})^{-\frac{n+1}{2}} \quad n = 1,2,\ldots \quad -\infty < x < \infty$$

siendo $\beta(p, q)$ la función beta:

$$\beta(p, q) = \frac{\Gamma(p).\Gamma(q)}{\Gamma(p+q)}$$

Al observar la distribución t de Student con n grados de libertad, observamos:

a) El campo de variabilidad de t_n es toda la recta real.

b) Su distribución depende sólo del parámetro n.

c) Conforme n va aumentado, la curva que representa a la función de densidad va siendo cada vez mas apuntada, coincidiendo en el límite, cuando n $\to\infty$ con la curva normal tipificada.

d) No depende de σ

e) Es simétrica respecto del eje de ordenadas OY.

f) En el muestreo al tomar muestras de tamaño n de media $\bar{x}$ y varianza s^2 de una población N(μ, σ), la variable

$$t_{n-1} = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n-1}}} = \frac{\bar{x}-\mu}{s}.\sqrt{n-1}$$

sigue una t de Student con n-1 grados de libertad. Esta propiedad es muy utilizada en la estimación y el contraste de hipótesis sobre la media de la población μ.

g) Está tabulada para diferentes valores de n. Para el uso de la tabla consideramos valores de $t_{\alpha,n}$ (punto crítico) tales que el área bajo la curva a la derecha de este valor es igual a α (nivel de significación). Es decir:

$$P(t \geq t_{\alpha;n}) = \alpha$$

sólo se encuentran valores de t ≥ 0 (o áreas α menores o iguales que 0.5). Para valores de t ≤ 0 se puede obtener por simetría.

[Volver al principio]

Distribución F de Fisher-Snedecor.

Sean χ_1^2 y χ_2^2 dos variables χ^2 de Pearson con n_1 y n_2 grados de libertad, independientes entre sí. Entonces la variable:

$$F_{n_1;n_2} = \frac{\dfrac{\chi_1^2}{n_1}}{\dfrac{\chi_2^2}{n_2}}$$

se le denomina F de Fisher_Snedecor con n_1 y n_2 grados de libertad. Tiene por función de densidad:

$$f_{n_1,n_2} = \begin{cases} \dfrac{\Gamma\left(\dfrac{n_1+n_2}{2}\right)\left(\dfrac{n_1}{n_2}\right)^{n_1/2}}{\Gamma\left(\dfrac{n_1}{2}\right)\Gamma\left(\dfrac{n_2}{2}\right)} \dfrac{x^{(n_1/2)-1}}{\left(1+\dfrac{n_1}{n_2}x\right)^{(n_1+n_2)/2}} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

En esta distribución hay que observar :

a) El campo de variabilidad de la F es el intervalo (0, +∞)

b) Depende de dos parámetros n_1 y n_2.

c) Es muy utilizada en análisis de la varianza.

d) Se encuentra tabulada para diferentes valores de n_1 y n_2. Para el uso de estas se consideran áreas a la derecha de un punto $F_{\alpha;n1,n2}$ (punto crítico) que deja a la derecha, bajo la curva un área igual a α (nivel de significación). $P(F_{n_1;n_2} \geq F_{\alpha;n_1;n_2}) = \alpha$.

En las tablas aparecen solo dos valores de α=0.05 y α=0.01. Para valores de α=0.95 y de α=0.99 se considera la relación:

$$F_{\alpha;n_1,n_2} = \frac{1}{F_{1-\alpha,n_1,n_2}}$$

[Volver al principio]

Distribución de Pareto

Una variable aleatoria X sigue una distribución de Pareto de parámetros α y x_0 si:

1. Puede tomar valores superiores o iguales a x_0.

2. Su función de densidad es:

$$f(x) = \begin{cases} \dfrac{\alpha}{x_0}\left(\dfrac{x_0}{x}\right)^{\alpha+1} & x \geq x_0 > 0; \quad \alpha > 0 \\ 0 & x < x_0 \end{cases}$$

En esta distribución hay que observar que:

a) Depende de los parámetros α y x_0.

b) Su función de distribución es:

$$F(x) = \begin{cases} 1 - \left(\dfrac{x_0}{x}\right)^{\alpha} & x \geq x_0 \\ 0 & x < x_0 \end{cases}$$

c) Su media y varianza para α >2 son:

$$\mu = \frac{\alpha x_0}{\alpha - 1}; \quad \sigma^2 = \frac{\alpha x^2{}_0}{(\alpha-1)^2(\alpha-2)}.$$

d) Es muy utilizada en economía, ya que la distribución de las rentas personales superiores a una cierta cantidad x_0 sigue una distribución de Pareto de parámetros α y x_0.

Se tiene

$$P(X > x) = 1 - P(X \leq x) = 1 - F(x) = \left(\frac{x_0}{x}\right)^{\alpha}$$

lo que nos permite determinar el número de personas que superan la renta x_0. El parámetro α se determina a partir de la media muestral. Normalmente toma valores próximos a 2.

[Volver al principio]

Distribución log-normal.

Una variable X sigue una distribución log-normal si los logaritmos neperianos de sus valores están normalmente distribuidos, es decir, si la variable Y=log X es N(μ,σ) es decir si:

$$f(x) = \begin{cases} \dfrac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(\log x-\mu)^2}{\sigma^2}} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

En ella observamos que:

a) Depende de los parámetros μ y σ.

b) Su media y varianza son:

$$\mu = e^{\mu+\sigma^2/2} \qquad \sigma^2 = e^{2\mu+2\sigma^2} - e^{2\mu+\sigma^2}$$

[Volver al principio]

La distribución Beta

Se ha utilizado para representar variables físicas cuyos valores se encuentran restringidos a un intervalo de longitud finita y para encontrar ciertas cantidades que se conocen como límites de tolerancia sin necesidad de la hipótesis de una distribución normal. Además la distribución beta juega un gran papel en la estadística bayesiana.

La v.a X sigue una distribución beta si su función de densidad está dada por:

$$f(x;\alpha,\beta) = \begin{cases} \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} & 0 < x < 1. \quad \alpha,\beta > 0 \\ 0 & otro \quad valor \end{cases}$$

Es fácil ver que si se reemplaza x por 1-x se tiene:

f(1-x; β, α)=f(x; α, β).

La función gamma.

Se define como:

$$\Gamma(n) = \int_0^\infty u^{n-1}.\exp(-u)du$$

Algunas propiedades de esta función son:

a) Si n es un entero positivo

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

$$b)\quad \Gamma(n+1) = n\Gamma(n);\quad n > 0$$

$$c)\quad \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

La función Beta.

La función beta se define como:

$$B(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx \quad siendo\ \alpha,\beta > 0$$

Puede demostrarse que la función beta y la gamma se encuentran relacionadas por la expresión:

$$B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

Media y varianza de la distribución beta

$$E[x] = \mu = \frac{\alpha}{\alpha+\beta} \qquad Var(x) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

[Volver al principio]

La distribución Gamma.

Supóngase que una pieza está sometida a una cierta fuerza, de manera que se romperá después de aplicar un número específico de ciclos fuerza. Si los ciclos ocurren de manera independiente y a una frecuencia promedio dada, entonces el tiempo que debe transcurrir antes de que el material se rompa es una variable aleatoria que sigue una distribución gamma. Su función de densidad viene expresada por:

$$f(x;\alpha,\theta) = \begin{cases} \dfrac{1}{\Gamma(\alpha)\theta^{\alpha}} x^{\alpha-1}.\exp(-\dfrac{x}{\theta}) & x > 0,\ \alpha,\theta > 0 \\ 0 & en\ otro\ caso \end{cases}$$

Esta distribución se emplea de manera extensa en una gran diversidad de áreas.

- Para representar el tiempo de falla de un sistema que falla solo si de manera exacta los componentes fallan y la falla de cada componente ocurre a una frecuencia constante $\lambda=1/\theta$ por unidad de tiempo.

- En líneas de espera para completar una reparación que se lleva a cabo en subestaciones; en cada una de las cuales es un evento independiente que ocurre a una frecuencia constante igual a $\lambda=1/\theta$.

- Ingresos familiares.

- Edad al contraer matrimonio.

Media y varianza de la distribución Gamma.

$$E[X]=\alpha.\theta \qquad Var[X]=\alpha\theta^2$$

[Volver al principio]

Distribución de Erlang.

La distribución gamma, cuando α es un entero positivo se conoce con el nombre de Erlang. Existe una asociación entre los modelos de probabilidad de Poisson y de Erlang. Si el número de eventos aleatorios independientes que ocurren en un lapso específico es una variable aleatoria de Poisson con frecuencia constante de ocurrencia igual a 1/ θ, entonces, para una α dada, el tiempo de espera hasta que ocurre el α-ésimo evento de Poisson sigue una distribución de Erlang.

Cuando α=1, la distribución de Erlang se reduce a una **distribución exponencial negativa**. Nótese que la variable aleatoria de una distribución exponencial negativa puede pensarse como el lapso que transcurre hasta el primer evento de Poisson. De acuerdo con esto, la variable aleatoria de Erlang es la suma de variables aleatorias independientes distribuidas exponencialmente.

Otro caso especial del modelo de probabilidad gamma es la **distribución chi-cuadrado**.

Si se hace $\alpha= \upsilon /2$ y $\theta=2$, se obtiene:

$$f(x;\nu)=\begin{cases}\dfrac{1}{\Gamma(\nu/2).2^{\nu/2}}.x^{\nu/2-1}.\exp(-x/2) & Si \quad x>0\\ 0 & en \quad otro \quad caso\end{cases}$$

donde υ recibe el nombre de grados de libertad.

La media y varianza de la distribución chi-cuadrado se obtienen de los de la gamma.

E[X]= υ y Var[X]=2. υ

[Volver al principio]

La distribución de Weibull.

El esfuerzo al que se someten los materiales puede modelarse por esta distribución. Se ha empleado mucho en situaciones del tipo tiempo-falla con el objetivo de lograr una amplia variedad de componentes mecánicos y eléctricos.

La función de densidad está dada por:

$$f(x;\alpha,\theta)=\begin{cases}\dfrac{\alpha}{\theta^{\alpha}}x^{\alpha-1}\exp\left[-\left(x/\theta\right)^{\alpha}\right] & x>0,\ \alpha,\theta>0\\ 0 & \textit{en otro caso}\end{cases}$$

Media y varianza de una distribución de Weibull.

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

Caso particulares

- Si $\alpha=1$ se reduce a la distribución exponencial negativa.
- Cuando $\alpha=2$ y $\theta=\sqrt{2}\,\sigma$.

$$f(x;\theta)=\begin{cases}\dfrac{1}{\theta}\exp(-\dfrac{x}{\theta}) & x>0,\ \theta>0\\ 0 & \end{cases}$$

que se conoce con el nombre de **distribución de Rayleigh.**

[Volver al principio]

Distribución exponencial negativa.

La distribución exponencial negativa puede emplearse para modelar el lapso de tiempo entre dos eventos consecutivos de Poisson que ocurren de manera independiente y a una frecuencia constante. Esta distribución se emplea con bastante frecuencia con objeto de modelar problemas del tipo tiempo-falla y como modelo para el intervalo en problemas de líneas de espera. Se demuestra que esta distribución no tiene “memoria”, en el sentido de que, la probabilidad de ocurrencia de eventos presentes o futuros no depende de los que hayan ocurrido en el pasado. De esta forma, la probabilidad de que una unidad falle en un lapso específico de tiempo depende nada mas que de la duración de este, no del tiempo en que la unidad ha estado en operación.

La función de densidad está dada por:

$$f(x;\theta) = \begin{cases} \frac{1}{\theta}\exp(-x/\theta) & x > 0, \theta > 0 \\ 0 & \end{cases}$$

Media y varianza de la distribución exponencial.

E[X]= θ; Var(X)= θ^2

[Volver al principio]