UNIDAD 3:
VARIABLES ALEATORIAS

Instructor: Fernando Sandoya, PhD.
Correo: f.sandoyas@.espol.edu.ec

ESTADÍSTICA
Fernando Sandoya, PhD.
Facultad de Ciencias Naturales y Matemáticas

Variables aleatorias

VARIABLES ALEATORIAS

Definición

Función que considerando los elementos del espacio muestral (Ω) puede tomar un valor definido dentro de los números reales.

$$X: \Omega \rightarrow \mathbb{R}$$

Es decir que para cada $\omega \in \Omega$ se tiene que la función $X(\omega)$ asigna un número real.

Ejemplo: Espacio muestral de los resultados del lanzamiento de 2 monedas

$$\Omega = \{cc, cs; sc; ss\}$$

Evento: Obtener cara

X: Número de caras obtenidas

¿Qué valores tomará X?

VARIABLES ALEATORIAS

Variable Aleatoria Discreta

Formalmente una v.a. **X** es discreta cuando el rango de X es un conjunto discreto (por ejemplo, el conjunto de números naturales). Intuitivamente, v.a.d. es aquella que toma valores que se pueden contar (números enteros),

Variable Aleatoria Continua

Formalmente una v.a. **X** es continua cuando el rango de X es un conjunto continuo (por ejemplo, un intervalo de números reales. Intuitivamente, es aquella que toma valores dentro de una escala continua (números reales). Sirven para medir

VARIABLES ALEATORIAS DISCRETAS

Función de probabilidad - Variables discretas

Sea X una v.a. discreta, se dice que la $f(x): \mathbb{R} \to \mathbb{R}$ es la función de distribución de probabilidad de X si cumple con:

$$\boldsymbol{f(x) \geq 0}$$

$$\sum_{\mathbf{x}} \mathbf{f(x) = 1}$$

$$\mathbf{P(X = x) = f(x)}$$

Si $A \subseteq \mathbb{R}$ entonces

$$\mathbf{P(X \in \mathit{A}) = \sum_{x \in \mathit{A}} f(x)}$$

VARIABLES ALEATORIAS DISCRETAS

Función Acumulada

Discretas

$F(x) = \mathbf{P(X \leq x)} = \sum_{t \leq x} \mathbf{f(t)}$; **para** $-\infty < x < \infty$

VARIABLES ALEATORIAS DISCRETAS

Momentos de variables Aleatorias Discretas

Media $$\boldsymbol{\mu = E(X) = \sum_{x} x\, f(x)}$$

E(X): se lee media de X, esperanza de la v.a. X o valor esperado de la v.a. X

Def: La esperanza de una función de la v.a.d. X, g(X) se define como:

$$\boldsymbol{E(g(X)) = \sum_{x} g(x)\, f(x)}$$

Def: Se define Varianza $\boldsymbol{\sigma^2 = V(X) = Var(X) = E(X-\mu)^2 = \sum_{x} (x-\mu)^2 f(x)}$

VARIABLES ALEATORIAS DISCRETAS

Esperanza - Propiedades

Sea X una v.a. con esperanza finita y c una constante. Entonces

a) $E(c) = c$

b) $E(cX) = cE(X)$

c) $Si\ X \geq 0\ entonces\ E(X) \geq 0$

d) $E(X+Y) = E(X) + E(Y)$

VARIABLES ALEATORIAS DISCRETAS

Varianza - Interpretación

Medida del grado de dispersión de los valores de la v.a. alrededor de la media, ponderados por sus respectivas probabilidades

Varianza – Propiedades

$a)\ V(X) \geq 0$

$b)\ V(c) = 0$

$c)\ V(cX) = c^2 V(X)$

$d)\ V(X + c) = V(X)$

$e)\ V(X + Y) \neq V(X) + V(Y), salvo\ el\ caso\ de\ independencia\ (se\ verá\ después)$

f) $V(X) = E(X^2) - \left(E(X)\right)^2 = E(X^2) - \mu^2$

VARIABLES ALEATORIAS DISCRETAS
Mas importantes

En la realidad aparecen con frecuencia las variables aleatorias discretas con diferentes funciones de distribución de probabilidad, a estas se las conoce como **Modelos Aleatorios Discretos**, y las principales son:

- **La distribución binomial**
- **La distribución binomial negativa (la distribución geométrica)**
- **La distribución hipergeométrica**
- **La distribución de Poisson**

VARIABLES ALEATORIAS DISCRETAS
Distribución Binomial

Antes de definir lo que es una variable aleatoria con distribución binomial, se debe definir el **experimento binomial.**

Experimento Binomial: Un experimento binomial es aquel que cumple lo siguiente:

1.- Consiste en n repeticiones o intentos similares.

2.- Cada repetición (o intento) toma uno de dos posibles resultados: éxito o fracaso

3.- La probabilidad éxito (representada con $\boldsymbol{p}$) se mantiene constante para cada repetición, al igual que la probabilidad de fracaso $q=(1-p)$

4.- Toda repetición es independiente de las demás

VARIABLES ALEATORIAS DISCRETAS
Distribución Binomial

Una variable aleatoria tiene distribución binomial con parámetros n, p: $X \sim B(n, p)$ si X "cuenta" el número de éxitos en n intentos en un experimento binomial. Donde p es la probabilidad de éxito en un intento.

Valores de X $\quad X: 0,1,2,3, \dots, n$

Función de probabilidad $\quad P(X = x) = \boldsymbol{p}(x) = \binom{n}{x} p^x q^{n-x}$

Parámetros $\quad n, p$

Media $\quad \mu = E(X) = np$

Varianza $\quad V(X) = npq$

VARIABLES ALEATORIAS DISCRETAS
Distribución Binomial

Ejemplo

Encuentre la probabilidad de que 7 personas de 10 se recuperarán de una enfermedad tropical si podemos suponer que la probabilidad que cualquiera de ellos se recuperará es de 0,40 (40%)

VARIABLES ALEATORIAS DISCRETAS
Distribución Binomial

Ejemplo

Con los datos del ejemplo anterior calcule la probabilidad que:

a) Menos de 3 personas se recuperen

b) 4 personas o más se recuperen

VARIABLES ALEATORIAS DISCRETAS
Distribución Binomial

En R:

Para cada distribución, R dispone de cuatro funciones. Se puede acceder a cada una de ellas simplemente precediendo el nombre de la distribución que figura en la tabla anterior por la letra que se indica a continuación:

- **d**: función de densidad o de probabilidad.
- **p**: función de distribución
- **q**: función para el cálculo de cuantiles.
- **r**: función para simular datos con dicha distribución.

Distribuciones Discretas

Distribución	Nombre en R
Binomial	binom
Poisson	pois
Geométrica	geom
Hipergeométrica	hyper
Binomial Negativa	nbinom

Distribución Binomial Negativa

En un experimento binomial, la variable que cuenta el número de intentos hasta lograr el k-ésimo éxito tiene distribución binomial negativa

Valores de X $X: k, k+1, k+2, k+3, \ldots$

Función de probabilidad $P(X = x) = f(x) = \binom{x-1}{k-1} p^k q^{x-k}$

Parámetros k, p

Media $E(X) = \frac{k}{p}$

Varianza $V(X) = \frac{kq}{p^2}$

VARIABLES ALEATORIAS DISCRETAS
Distribución Binomial Negativa

<u>Ejemplo</u>

Se está realizando un estudio de la tasa de éxito de un nuevo antibiótico contra una bacteria específica. Supongamos que un investigador desea determinar cuántos intentos de cultivo son necesarios para obtener tres cultivos exitosos donde la bacteria es efectivamente inhibida por el antibiótico. Los 3 cultivos exitosos son exigidos por una norma de calidad.

El investigador sabe, basado en estudios previos, que la probabilidad de éxito de inhibir la bacteria con una sola aplicación del antibiótico es del 20%. El investigador desea saber la probabilidad de que se necesiten exactamente 10 intentos de cultivo para lograr 3 éxitos (es decir, 3 cultivos donde la bacteria es inhibida por el antibiótico).

VARIABLES ALEATORIAS DISCRETAS
Distribución Geométrica

Si X tiene distribución binomial negativa con ***K*=1**, se dice que tiene distribución Geométrica.

Es decir, se dice que X tiene distribución geométrica con parámetro *p* si X cuenta el número de intentos hasta obtener el primer éxito.

Por tanto, la función de distribución de probabilidad de una variable geométrica es:

$$\boldsymbol{f(x) = p\,(1-p)^{x-1}, x = 1, 2, 3, \ldots}$$

VARIABLES ALEATORIAS DISCRETAS
Distribución Geométrica

Una empresa de ingeniería está probando la eficiencia de una máquina para detectar defectos en piezas de un lote de producción. Se conoce que el 15% de las piezas producidas por la máquina son defectuosas, y las pruebas se realizan de una en una, revisando una pieza tras otra hasta encontrar la primera pieza defectuosa.

Preguntas:

a- ¿Cuál es la probabilidad de que la primera pieza defectuosa se encuentre en la quinta inspección?

b- Si el inspector decide revisar 10 piezas, ¿cuál es la probabilidad de que no se haya encontrado ninguna pieza defectuosa después de estas 10 inspecciones?

VARIABLES ALEATORIAS DISCRETAS
Distribución Hipergeométrica

Distribución Hipergeométrica

Número de éxitos en una muestra de *n* elementos tomados de una población *N* de los cuáles *k* elementos tienen una característica de interés

Valores de X $X: Max\{0, K-(N-n)\}, \dots, Min\{n, K\}$

Función de probabilidad $$P(X=x)=f(x)=\frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$$

Parámetros N, n, k

Media $$E(X)=\frac{nk}{N}$$

Varianza $$V(X)=\frac{N-n}{N-1}\left(\frac{nk}{N}\right)\left(1-\frac{k}{N}\right)$$

VARIABLES ALEATORIAS DISCRETAS
Distribución Hipergeométrica

Ejemplo

Se está realizando un estudio de la distribución de bacterias resistentes a antibióticos en una población de bacterias. Supongamos que un laboratorio tiene una colonia de 1000 bacterias, de las cuales se sabe que 150 son resistentes a un cierto antibiótico. Se selecciona una muestra aleatoria de 100 bacterias para un estudio más detallado.

Pregunta:

1. ¿Cuál es la probabilidad de que exactamente 10 bacterias en la muestra sean resistentes al antibiótico?
2. Si aproximáramos con una distribución binomial, ¿cuál sería la probabilidad?
3. ¿Cuál es el cálculo correcto?

VARIABLES ALEATORIAS DISCRETAS
Distribución de Poisson

Distribución Poisson $X \sim P(\lambda), \lambda > 0$

Una variable de Poisson con promedio λ cuenta el Número de eventos que ocurren en una unidad (de tiempo, de área, de volumen o región)

Valores de X $X: 0,1,2,3, \ldots$

Función de probabilidad $$P(X = x) = f(x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

Parámetros λ

Media $E(X) = \lambda$ **Varianza** $V(X) = \lambda$

VARIABLES ALEATORIAS DISCRETAS
Distribución de Poisson

Ejemplo:

Un equipo de microbiólogos ambientales está llevando a cabo un estudio para evaluar la presencia de microorganismos patógenos en un río. Se ha observado que, en promedio, hay 2 microorganismos patógenos por hora en el agua del río. Calcular la probabilidad de observar hasta 3 microorganismos patógenos en una muestra de agua tomada del río durante una hora.

VARIABLES ALEATORIAS CONTINUAS

Función de probabilidad - Variables continuas

Sea X una v.a. continua, se dice que la $f(x)$: $\mathbb{R} \rightarrow \mathbb{R}$ es la función de densidad de X si cumple con:

$i.\ f(x) \geq 0$

$ii.\ \int_{-\infty}^{+\infty} f(x)dx = 1$

Y la probabilidad de que X tome valores en un intervalo se calcula como:

$$P(a < X < b) = \int_{a}^{b} f(x)dx$$

VARIABLES ALEATORIAS CONTINUAS

Función Acumulada

Continuas

$$F(x) = \mathbf{P(X \leq x)} = \int_{-\infty}^{x} f(t)dt \;\; ; \;\; \textbf{para} -\infty < \boldsymbol{x} < \infty$$

Prop:

1. f(x) = d(F(x))/dx

2. F(x) es una función creciente

3. F(x) es una función contínua

Distribución Uniforme

Probabilidad es uniforme en un intervalo cerrado [a,b]

Función de probabilidad

$$f(x) = \begin{cases} \dfrac{1}{b-a}; & a \leq X \leq b \\ 0 & en\ cualquier\ otro\ caso \end{cases}$$

Parámetros a, b

Media $\mathrm{E(X)} = \dfrac{\mathrm{a+b}}{2}$

Varianza $\mathrm{V(X)} = \dfrac{\mathrm{(b-a)^2}}{12}$

VARIABLES ALEATORIAS CONTINUAS
Distribución Uniforme

Ejemplo

Se está realizando un estudio del tiempo que tardan diferentes cepas de bacterias en metabolizar un sustrato específico bajo condiciones controladas. Supongamos que, basado en conocimientos previos y experimentos preliminares, un investigador determina que el tiempo necesario para que una cepa bacteriana particular metabolice completamente el sustrato varía uniformemente entre 2 y 6 horas.

El investigador está interesado en calcular la probabilidad de que la bacteria metabolice el sustrato en un tiempo controlable, y se ha establecido que si la bacteria completa el metabolismo del sustrato entre 3 y 4 horas este proceso es controlable. Calcular la probabilidad de que este proceso sea controlable.

VARIABLES ALEATORIAS CONTINUAS
Distribución Normal

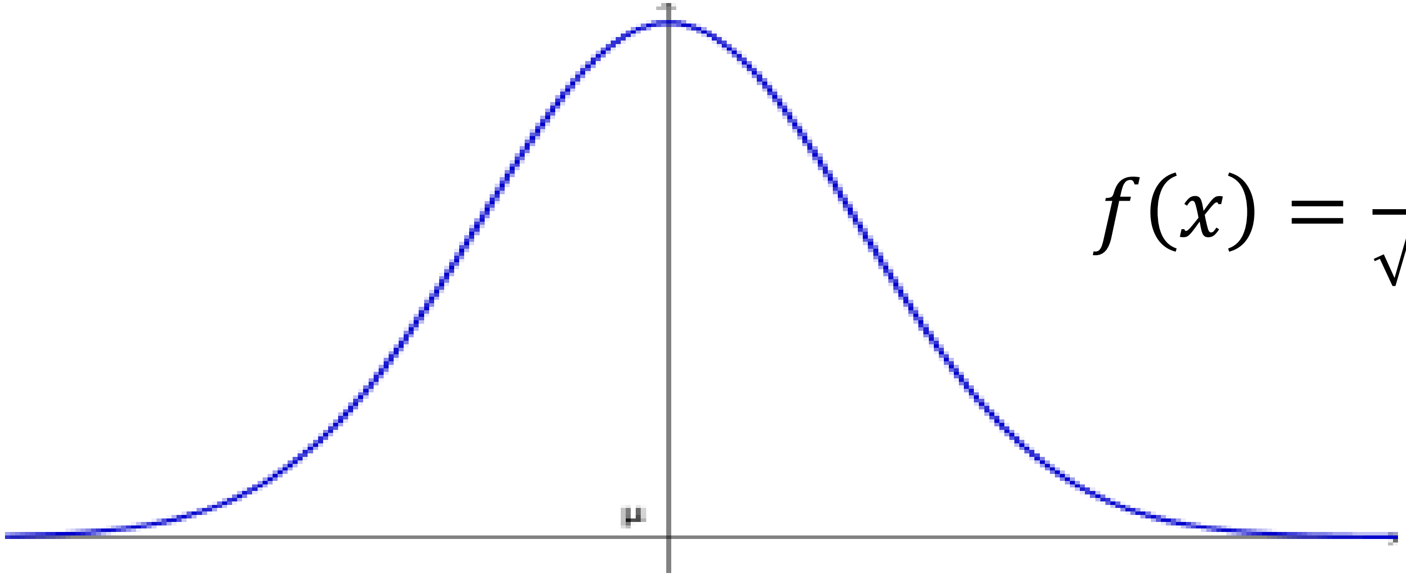
Distribución Normal: Es la piedra angular de la teoría estadística moderna. Conocida y estudiada desde hace mucho tiempo, es utilizada para describir el comportamiento aleatorio de muchos procesos que ocurren en la naturaleza y también realizados por los humanos

- Muchos fenómenos de la naturaleza siguen está distribución
- Se identifican los parámetros μ para la media y σ^2 para la varianza
- La curva es simétrica en μ
- La curva es asintótica al eje horizontal conforme se aleja de μ en cualquier dirección
- El área bajo la curva es 1

Distribución Normal

Parámetros: μ; σ^2

Función de probabilidad



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

VARIABLES ALEATORIAS CONTINUAS
Distribución Normal

Ejemplo

Se está realizando un estudio de la concentración de una toxina producida por una cepa bacteriana en condiciones específicas de cultivo. Supongamos que, después de realizar numerosos experimentos, se determina que la concentración de esta toxina en el medio de cultivo sigue una distribución normal con una media de 10 μg/ml y una desviación estándar de 2 μg/ml.

El equipo de investigación desea saber la probabilidad de que, en un experimento aleatorio, la concentración de toxina en el medio de cultivo esté entre 8 μg/ml y 12 μg/ml, lo cual es crucial para evaluar el riesgo potencial para la salud y para diseñar estrategias de mitigación adecuadas.

VARIABLES ALEATORIAS CONTINUAS
Distribución Normal

Ejemplo (inverso de la distribución normal)

un investigador está estudiando la sensibilidad de una cepa bacteriana a un nuevo antibiótico. A través de varios experimentos, determina que la concentración mínima inhibitoria (CMI), que es la concentración más baja de antibiótico necesaria para inhibir visiblemente el crecimiento de la bacteria, sigue una distribución normal. Basado en los datos recopilados, la CMI tiene una media de 5 μg/ml y una desviación estándar de 1 μg/ml.

El investigador quiere determinar la concentración de antibiótico necesaria para inhibir el crecimiento del 95% de la población bacteriana. Este valor es crucial para establecer dosis efectivas y seguras del antibiótico.

VARIABLES ALEATORIAS CONTINUAS
Distribución Normal

<u>Definición:</u> Si una variable tiene una distribución Normal, mediante una sustitución se la puede transformar a otra variable con una Distribución Normal Estándar. Este cambio facilita el cálculo de probabilidad y se denomina estandarización.

Notación

$X \sim N(\mu, \sigma)$ Define a **X** como una variable con **distribución Normal** con media μ y desviación estándar σ

$Z \sim N(0, 1)$ Define a **Z** como una variable con **distribución Normal** Estándar con media **0** y desviación estándar **1**

Sea **X** una variable aleatoria con distribución Normal: $X \sim N(\mu, \sigma)$,

Entonces, la variable aleatoria $Z = \dfrac{X - \mu}{\sigma}$

Tiene distribución Normal Estándar: $Z \sim N(0, 1)$

VARIABLES ALEATORIAS CONTINUAS
Distribución Normal

Distribución Normal Estándar

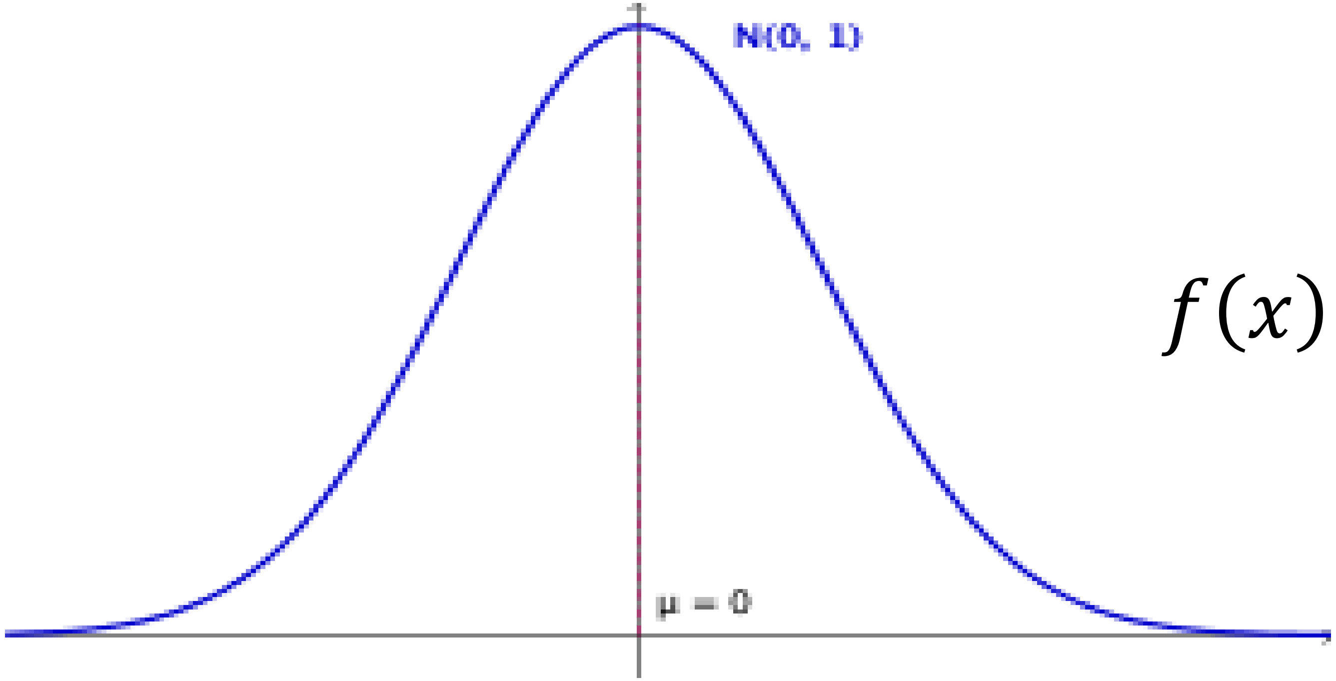
Valores de X: - ∞ ; +∞

$$Z = \frac{(x-\mu)}{\sigma}$$

Media $\boldsymbol{E(X) = 0}$

Varianza $\boldsymbol{V(X) = 1}$

Función de probabilidad

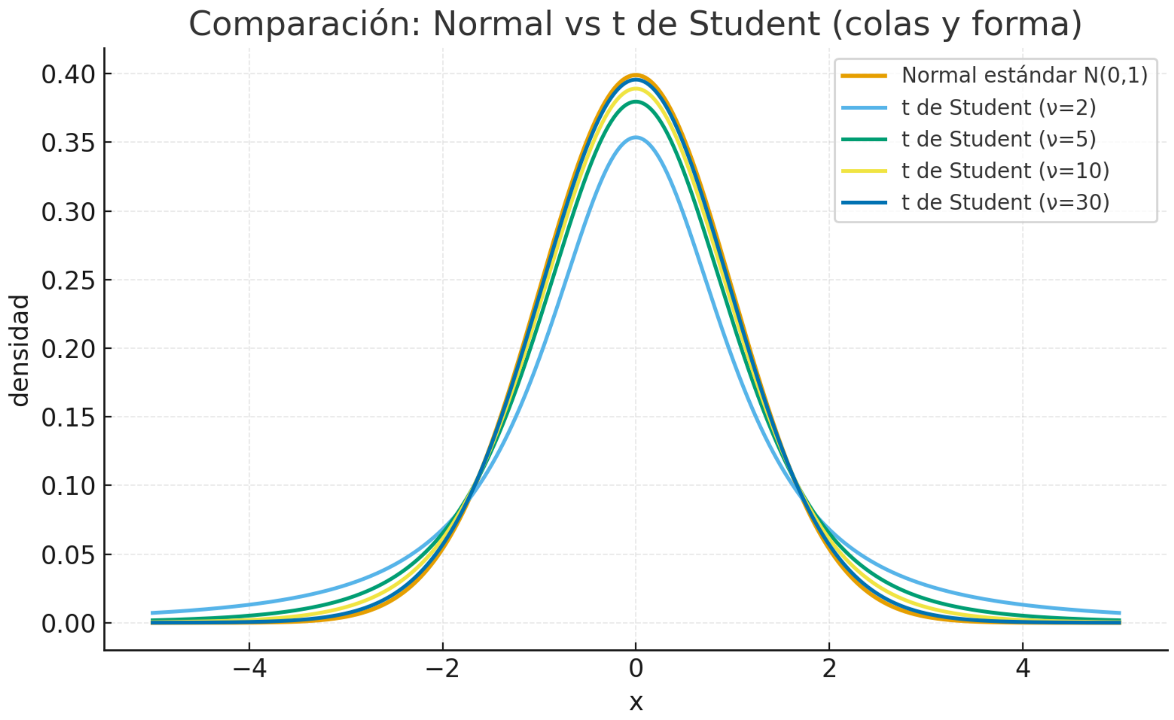


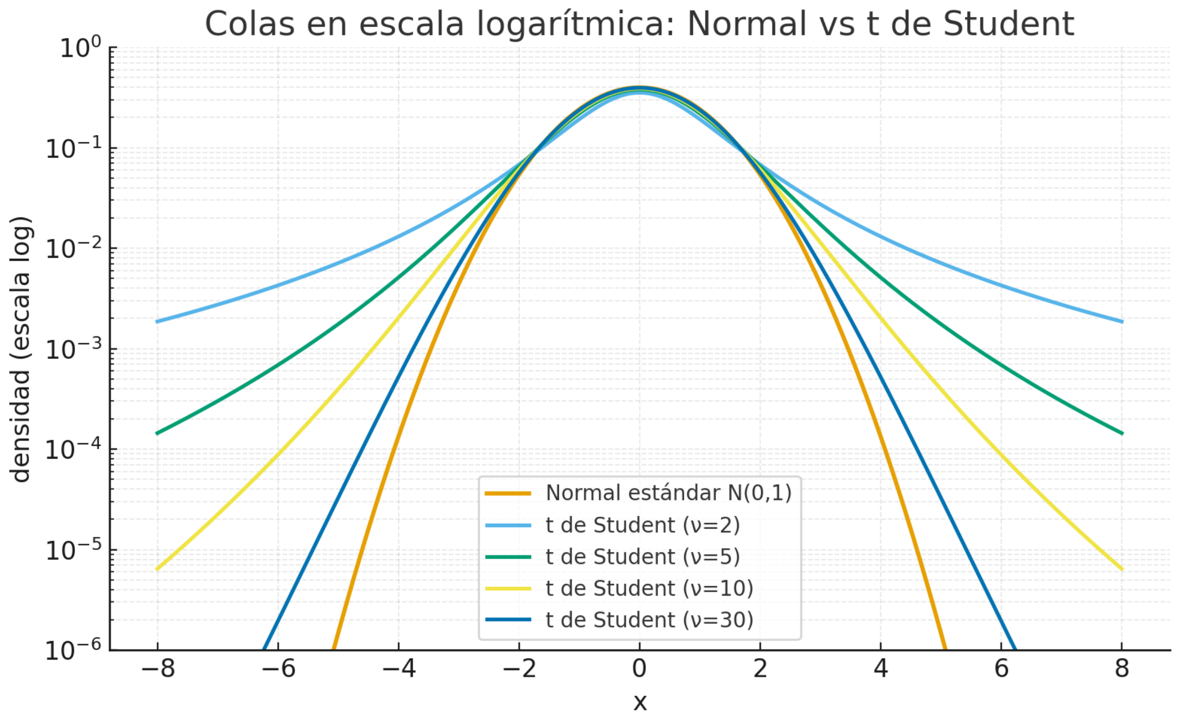
$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

VARIABLES ALEATORIAS CONTINUAS
Distribución t-student

- **Definición:** La distribución t-Student es una distribución continua simétrica con cola más pesada que la Normal.
- **Parámetro:** grados de libertad (gl = v). La forma depende de v. Se representa con t_v
- **Relación con la Normal:** Si v es grande la distribución t-Student se aproxima a la distribución normal: $v \to \infty$, $t_v \to N(0,1)$.
- **Simetría:** centrada en 0.

VARIABLES ALEATORIAS CONTINUAS
Distribución t-student





Distribución Exponencial

X tiene una distribución exponencial, con parámetro β > 0; si su función de densidad está dada por

$$f(x) = \begin{cases} \dfrac{1}{\beta} e^{-\frac{x}{\beta}} & x > 0 \\ 0 & Si\ no \end{cases}$$

Media $\mathrm{E}(\mathrm{X}) = \beta$ **Varianza** $\mathrm{V}(\mathrm{X}) = \beta^2$

VARIABLES ALEATORIAS CONTINUAS
Distribución Beta

Definición: Tiene aplicaciones importantes por la variedad de formas diferentes que puede tomar su función de densidad eligiendo valores para sus parámetros.

Una variable aleatoria continua **X** tiene **distribución Beta** si su densidad de probabilidad está dada por

$$f(x) = \begin{cases} \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, & 0 < x < 1 \\ 0, & \text{para otro } x \end{cases}$$

En donde $\alpha>0$, $\beta>0$ son los parámetros para este modelo. $\Gamma(\,)$ es la función Gamma

Si n es natural $\Gamma(n)=(n-1)!$

VARIABLES ALEATORIAS DISCRETAS
Distribución Beta

Ejemplo

Se está realizando un estudio sobre la tasa de supervivencia de una bacteria específica en respuesta a diferentes concentraciones de un desinfectante. El equipo de investigación está estudiando cómo la concentración de un nuevo desinfectante afecta la tasa de supervivencia de una cepa bacteriana particular. Después de realizar varios experimentos preliminares, el equipo tiene algunas expectativas sobre la tasa de supervivencia de las bacterias: creen que, en una concentración específica del desinfectante, la tasa de supervivencia es relativamente baja, pero hay cierta incertidumbre debido a la variabilidad en los datos experimentales.
Para modelar esta incertidumbre y actualizar sus creencias con más datos, deciden utilizar la distribución beta con $\alpha=2$ y $\beta=5$ para reflejar su creencia de que la tasa de supervivencia es baja a una cierta concentración del desinfectante, pero con cierta incertidumbre. Se quiere calcular la probabilidad de que la tasa de supervivencia real de las bacterias sea mayor que 20%, lo cual sería un riesgo peligroso.

Distribución Exponencial

X tiene una distribución exponencial, con parámetro β > 0; si su función de densidad está dada por

$$f(x) = \begin{cases} \frac{1}{\beta} e^{-\frac{x}{\beta}} & x > 0 \\ 0 & Si\ no \end{cases}$$

Media $\mathrm{E(X)} = \beta$

Varianza $\mathrm{V(X)} = \beta^2$

VARIABLES ALEATORIAS CONTINUAS
Distribución Exponencial

Distribución Exponencial (λ)

Cuando el número de eventos por unidad de tiempo tiene una distribución de Poisson con parámetro λ entonces el tiempo entre eventos tiene distribución exponencial con tasa λ

$$f(t) = \begin{cases} \lambda e^{-\lambda t} & t > 0 \\ 0 & Si\ no \end{cases}$$

$$\mathrm{F}(t) = \begin{cases} 1 - e^{-\lambda t} & t \geq 0 \\ 0 & t < 0 \end{cases}$$

Media $\mathrm{E(T)} = \dfrac{1}{\lambda}$

Varianza $\mathrm{V(T)} = \dfrac{1}{\lambda^2}$

VARIABLES ALEATORIAS CONTINUAS
Distribución Exponencial

Ejemplo

un equipo de investigación desea estudiar la efectividad de un nuevo antibiótico. En particular, están interesados en el tiempo que tardan las bacterias en morir después de la exposición al antibiótico. Los estudios preliminares sugieren que este tiempo de supervivencia sigue una distribución exponencial. A través de experimentos controlados, el equipo determina que la tasa promedio de muerte de las bacterias bajo la influencia de este antibiótico es de 0.25 por hora, lo que significa que en promedio, las bacterias sobreviven 4 horas después de la exposición al antibiótico.

a. Determinar la probabilidad que las bacterias sobrevivan menos de 3 horas luego de exponerlas al antibiótico.

Distribución Gamma

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

Si $\quad u = x^{\alpha-1} \quad$ y $\quad dv = e^{-x} dx \qquad \Gamma(\alpha) = (\alpha-1)\int_0^{\infty} x^{\alpha-2} e^{-x} dx$

Es decir $\qquad \Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$

$$\Gamma(\alpha) = (\alpha-1)(\alpha-2)\Gamma(\alpha-2)$$

Si α es un entero positivo n $\boldsymbol{\Gamma(n) = (n-1)!}$

Si α = 1/2 $\qquad \boldsymbol{\Gamma(1/2) = \sqrt{\pi}}$

Distribución Gamma

X tiene una distribución Gamma, con parámetros α > 0 y β > 0; si su función de densidad está dada por

$$f(x) = \begin{cases} \dfrac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} & x > 0 \\ 0 & Si\ no \end{cases}$$

Media $\mathrm{E}(\mathrm{X}) = \alpha\beta$

Varianza $\mathrm{V}(\mathrm{X}) = \alpha\beta^2$

VARIABLES ALEATORIAS CONTINUAS
Distribución Chi-cuadrado χ^2

Distribución Xi - Cuadrado

Si una distribución Gamma, tiene parámetros α = *v/2* y β = 2; su función de densidad se denomina Xi-Cuadrado y está dada por

$$f(x) = \begin{cases} \dfrac{1}{2^{\frac{v}{2}}\Gamma\left(\frac{v}{2}\right)} x^{\frac{v}{2}-1} e^{-\frac{x}{2}} & x > 0; v\ es\ entero\ positivo \\ 0 & Si\ no \end{cases}$$

Media $\mathrm{E(X) = v}$

Varianza $\mathrm{V(X) = 2v}$

VARIABLES ALEATORIAS CONTINUAS
Distribución Weibull

Definición: Se usa en problemas relacionados con fallas de materiales y estudios de confiabilidad

Una variable aleatoria continua **X** tiene **distribución de Weibull** si su densidad de probabilidad está dada por

$$f(x) = \begin{cases} \alpha\beta x^{\beta-1}e^{-\alpha x^{\beta}}, & x > 0 \\ 0, & \text{para otro } x \end{cases}$$

En donde $\alpha>0$, $\beta>0$ son los parámetros para este modelo

DISTRIBUCIONES CONJUNTAS

Distribuciones Conjuntas
CASO DISCRETO:

Sean X e Y variables aleatorias DISCRETAS se dice que la función f(x,y) es una distribución de probabilidad conjunta si:

- $f(x,y) \geq 0;\ \forall\ x,y$
- $\sum_x \sum_y f(x,y) = 1$
- $P(X = x; Y = y) = f(x,y)$
- $Si\ x,y\ \in A\ P\big((X,Y) \in A\big) = \sum_A \sum f(x,y);$

DISTRIBUCIONES CONJUNTAS

Distribuciones Conjuntas
CASO CONTINUO:

Sean X e Y variables aleatorias CONTINUAS se dice que la función f(x,y) es una distribución de densidad conjunta si:

- $f(x,y) \geq 0;\ \forall\, x, y$
- $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dxdy = 1$
- $P\big((X,Y) \in A\big) = \iint_A f(x,y)dxdy\,;\, siendo\ A\ una\ región\ en\ el\ plano\ xy$

DISTRIBUCIONES CONJUNTAS

Distribuciones Conjuntas - Marginales

Sea la función f(x,y) la función de distribución de probabilidad/densidad conjunta de las variables aleatorias X e Y; sus correspondientes marginales serán:

$$f_X(\mathrm{x}) = \sum_{\mathrm{y}} \mathrm{f(x,y)} \qquad f_Y(\mathrm{y}) = \sum_{\mathrm{x}} \mathrm{f(x,y)}$$

$$f_X(\mathrm{x}) = \int_{\mathrm{y}} \mathrm{f(x,y)dy} \qquad f_Y(\mathrm{y}) = \int_{\mathrm{x}} \mathrm{f(x,y)dx}$$

DISTRIBUCIONES CONJUNTAS

Distribuciones Conjuntas - Independencia

Sea la función f(x,y) la función de distribución de probabilidad/densidad conjunta de las variables aleatorias X e Y; se dice que X e Y son estadísticamente independientes si y sólo si se cumple:

$$f(x,y) = f_X(x) f_Y(y);\ \forall x, \forall y$$

DISTRIBUCIONES CONJUNTAS

Distribuciones Conjuntas - Esperanza

Sea la función f(x,y) la función de distribución de probabilidad/densidad conjunta de las variables aleatorias X e Y; y g(x,y) una función cualquiera de las variables aleatorias X, Y, se define la Esperanza de g(x,y) como:

$$E(g(x,y)) = \sum_x \sum_y g(x,y) f(x,y)$$

$$E(g(x,y)) = \int_y \int_x g(x,y) f(x,y)\, dxdy$$

DISTRIBUCIONES CONJUNTAS

Distribuciones Conjuntas - Esperanza

En particular se define la esperanza conjunta de la siguiente manera:

$$E(xy) = \sum_{x} \sum_{y} xyf(x, y)$$

$$E(xy) = \int_{y} \int_{x} xyf(x, y)\, dxdy$$

DISTRIBUCIONES CONJUNTAS

Distribuciones Conjuntas – Covarianza

Sea la función f(x,y) la función de distribución de probabilidad/densidad conjunta de las variables aleatorias X e Y; se define la covarianza entre las variables de la siguiente manera:

$$\sigma_{X,Y}^2 = COV(X,Y) = E\left[(x-\mu_x)(y-\mu_y)\right] = \sum_x \sum_y (x-\mu_x)(y-\mu_y) f(x,y)$$

$$\sigma_{X,Y}^2 = COV(X,Y) = E\left[(x-\mu_x)(y-\mu_y)\right] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (x-\mu_x)(y-\mu_y) f(x,y)\, dxdy$$

PROPIEDAD: $\sigma_{X,Y}^2 = COV(X,Y) = E[X\,Y] - E[X]E[Y]$

DISTRIBUCIONES CONJUNTAS

Distribuciones Conjuntas – Covarianza

propiedades

Si X, Y, W, y V son variables aleatorias y a, b, c, d son constantes ("constante" en este contexto significa no aleatorio), se cumple que:

1. $\mathrm{Cov}(X, a) = 0$
2. $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$, la varianza de X
3. $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$
4. $\mathrm{Cov}(aX, bY) = ab\ \mathrm{Cov}(X, Y)$
5. $\mathrm{Cov}(X + a, Y + b) = \mathrm{Cov}(X, Y)$
6. $\mathrm{Cov}(aX + bY, cW + dV) = ac\ \mathrm{Cov}(X, W) + ad\ \mathrm{Cov}(X, V) + bc\ \mathrm{Cov}(Y, W) + bd\ \mathrm{Cov}(Y, V)$
7. $\mathrm{Cov}(X, Y) = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y)$ fórmula que suele emplearse en la práctica para calcular la covarianza.

Estas propiedades se deducen de manera casi directa de la definición de la covarianza. La covarianza trata de explicar qué tan relacionadas se encuentran dos variables entre sí, qué tanto se mueve una cuando la otra se mueve otro tanto. Ejemplo, si la variable X se mueve 1, supongamos que la variable Y se mueve 2, entonces podemos decir que la variable Y se mueve positivamente el doble de lo que se movería la variable X.

Distribuciones Conjuntas: Coeficiente de Correlación

Sea la función f(x,y) la función de distribución de probabilidad/densidad conjunta de las variables aleatorias X e Y; se define el **coeficiente de correlación lineal** entre las variables X e Y de la siguiente manera:

$$\rho_{x,y} = \frac{\sigma_{X,Y}^2}{\sigma_x \sigma_y}$$

Donde: $\sigma_X = \sqrt{\sigma_X^2}$; $\sigma_Y = \sqrt{\sigma_Y^2}$