



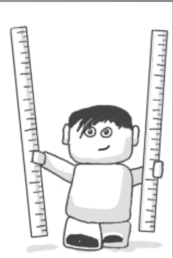
Teoría de probabilidades III

Curso 2023-24

Facultad de Física

Técnicas Experimentales II

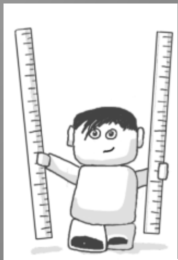
USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA



Objetivos tema:

- **Transformación de la variable aleatoria. Transformación de la función de probabilidad y medidas características.**
- **Distribución de probabilidad de Bernoulli.**
- **Distribución binomial.**
- **Ley de los grandes números.**
- **Distribución geométrica o de Pascal.**
- **Distribución de Poisson.**
- **Distribución normal o de Gauss.**

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA



Transformación de la variable aleatoria

En el caso de una variable aleatoria discreta *X* tendremos la posibilidad de redefinirla mediante una aplicación *g(x)* de $\mathbb{R}$ en $\mathbb{R}$

$$\begin{array}{ccccc} \Omega & \xrightarrow{X} & \mathbb{R} & \xrightarrow{g} & \mathbb{R} \\ A_i & \hookrightarrow & x_i & \hookrightarrow & y_i = g(x_i) \end{array}$$

Si la función *g(x)* es uno a uno, entonces podremos escribir

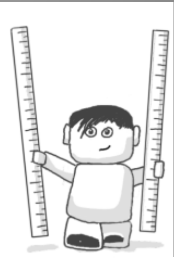
$$p_Y(y_i) = p_X(x_i)$$

En general definiremos el conjunto

$$I_j = \{i \in X(\Omega) |\, g(x_i) = y_j\}$$

$$p_Y(y_j) = \sum_{i \in I_j} p_X(x_i)$$

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

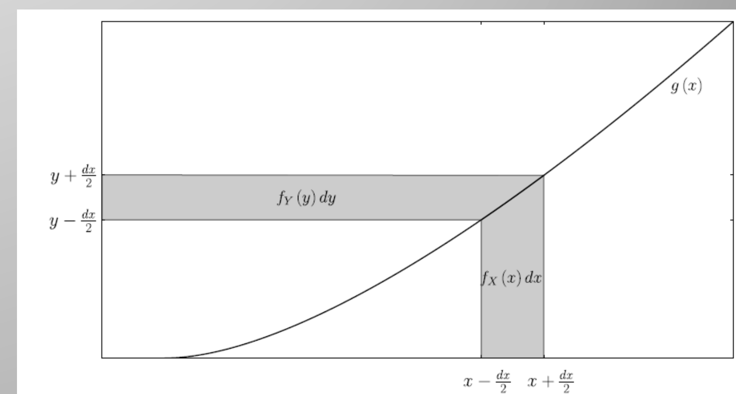


Transformación de la variable aleatoria

En el caso de una variable aleatoria continua *X* redefinida mediante una aplicación *g(x)* de $\mathbb{R}$ en $\mathbb{R}$

$$
\begin{array}{ccccc}
\Omega & \xrightarrow{\quad X \quad} & \mathbb{R} & \xrightarrow{\quad g \quad} & \mathbb{R} \\
A & \hookrightarrow & x & \hookrightarrow & y = g(x)
\end{array}
$$

Sean $\boldsymbol{f_X(x)}$ $\boldsymbol{f_Y(y)}$ las funciones de densidad de probabilidad de las variables aleatorias *X* e *Y*

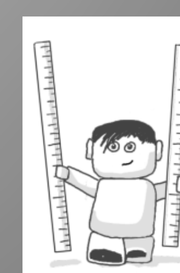


$$dp\big(y = g(x)\big) = |\, f_Y(y)\, dy \,| = |\, f_X(x)\, dx \,|$$

Por lo tanto:

$$f_Y(y) = \frac{f_X(x)}{\left|\frac{dy}{dx}\right|}$$

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA



Reusando variables aleatorias

En muchos ordenadores encontramos generadores aleatorios de distribución de probabilidad uniforme. Usando el resultado anterior veamos cómo construir variables aleatorias con nuevas distribuciones.

Supongamos que X es una variable aleatoria uniforme entre 0 y 1 y queremos construir una variable aleatoria Y entre 0 y 1 con densidad de probabilidad y^2.

$$f_X(x) = 1 \qquad 0 \leq x \leq 1$$

$$f_Y(y) = k\, y^2 \qquad 0 \leq y \leq 1$$

Como:

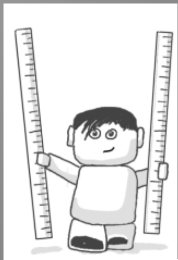
$$f_Y(y) = \frac{f_X(x)}{\left|\frac{dy}{dx}\right|}$$

$$\left|\frac{dy}{dx}\right| = \frac{1}{y^2}$$

$$y^2 dy = dx$$

$$y = \sqrt[3]{x}$$

$$f_Y(y) = \frac{1}{\left|\frac{dy}{dx}\right|} = \frac{1}{\frac{1}{3}\frac{1}{y^2}} = 3\, y^2$$



Revisitando la distribución exponencial

En el cálculo de Monte Carlo en la interacción radiación materia muchos procesos tienen una probabilidad de interacción exponencial en función de la profundidad

$$f_X(x) = 1$$

$$f_Y(y) = \beta\, e^{-\beta y} \qquad 0 \leq y < +\infty$$

Como:

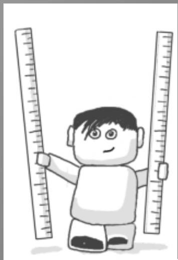
$$f_Y(y) = \frac{f_X(x)}{\left|\frac{dy}{dx}\right|}$$

$$\left|\frac{dy}{dx}\right| = \frac{1}{\beta\, e^{-\beta y}}$$

$$\frac{1}{\beta}\left(\beta - \beta\, e^{-\beta y}\right) = x$$

$$y = -\frac{1}{\beta}\log(1-x)$$

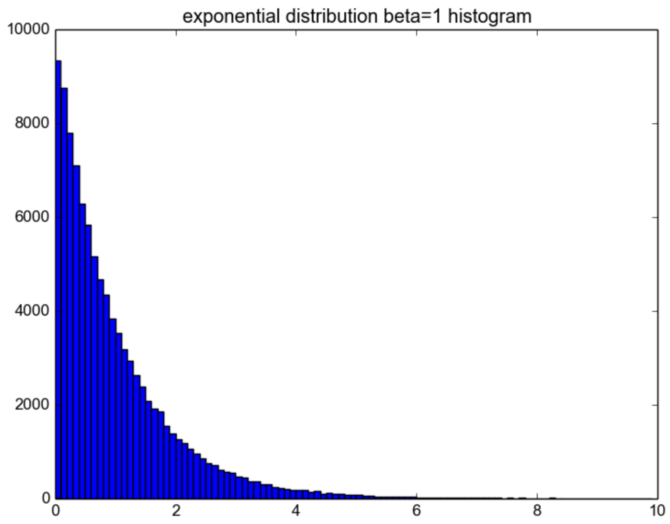
$$0 \leq x < 1$$



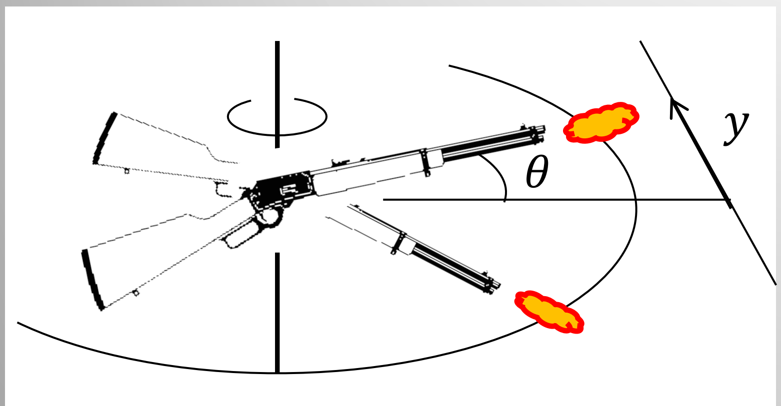
Distribución exponencial en python

Distribución exponencial $\boldsymbol{\beta = 1/scale}$

```
import numpy as np
from matplotlib import pyplot as plt
x=np.random.uniform(size=100000)
beta=1
y=-1/beta*np.log(1-x)
bins=np.arange(0,10,0.1)
plt.hist(y,bins)
plt.title(“exponential distribution beta=1 histogram”)
plt.show()
```



Disparando al azar



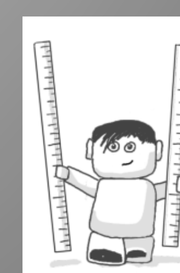
Supongamos una escopeta que gira sobre un eje disparando al azar en un plano. La distribución de probabilidad en ángulo es uniforme:

$$f_\theta(\theta) = \frac{1}{\pi} \qquad -\frac{\pi}{2} \leq x \leq \frac{\pi}{2}$$

$$y = g(\theta) = \mathrm{tg}(\theta)$$

$$f_Y(y) = \frac{f_\theta(\theta)}{\left|\frac{dy}{d\theta}\right|} = \frac{1}{\pi\,(1+\mathrm{tg}^2(\theta))} = \frac{1}{\pi\,(1+y^2)} \qquad -\infty < y < +\infty$$

Esta es la denominada distribución de Cauchy que aunque está bien normalizada y tiene media, por simetría, nula; su varianza y todos los momentos superiores a orden dos son divergentes!!



Media y varianza

En general también nos interesa saber cómo se transforma la media y la varianza de las variables aleatorias en un cambio de variable.

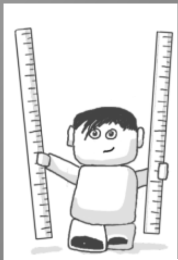
$$y = g(x)$$

$$g(x) = g(\mu_X) + \left.\frac{dg}{dx}\right|_{x=\mu_X} (x-\mu_X) + \frac{1}{2}\left.\frac{d^2g}{dx^2}\right|_{x=\mu_X} (x-\mu_X)^2 + \ \ldots$$

$$\mathbb{E}\{y\} = \mathbb{E}\{g(x)\} = g(\mu_X) + \left.\frac{dg}{dx}\right|_{x=\mu_X} \mathbb{E}\{(x-\mu_X)\} \ + \frac{1}{2}\left.\frac{d^2g}{dx^2}\right|_{x=\mu_X} \mathbb{E}\{(x-\mu_X)^2\} + \ \ldots$$

En aquellos casos en que podramos despreciar los términos de orden superior (momentos y derivadas de orden superior pequeños, datos muy agrupados entorno a la media),

$$\mathbb{E}\{y\} \approx g(\mu_X)$$



Media y varianza

Considerando válida la aproximación a primer orden

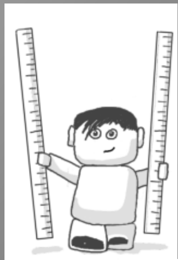
$$g(x) \approx g(\mu_X) + \left.\frac{dg}{dx}\right|_{x=\mu_X} (x-\mu_X)$$

$$\sigma^2{}_Y = \mathbb{E}\{(y-\mu_Y)^2\} \approx \mathbb{E}\left\{\left(g(\mu_X) + \left.\frac{dg}{dx}\right|_{x=\mu_X} (x-\mu_X) - \mu_Y\right)^2\right\}$$

$$\sigma^2{}_Y \approx \left(\left.\frac{dg}{dx}\right|_{x=\mu_X}\right)^2 \mathbb{E}\{(x-\mu_X)^2\}$$

$$\sigma^2{}_Y \approx \left(\left.\frac{dg}{dx}\right|_{x=\mu_X}\right)^2 \sigma^2{}_X$$

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA



Ejemplo

Hemos visto en la transformación de variable aleatoria:

$f_X(x) = 1 \qquad 0 \leq x \leq 1$

$f_Y(y) = 3\, y^2 \qquad 0 \leq y \leq 1$

$y = \sqrt[3]{x}$

$$\sigma^2{}_Y \approx \left(\left. \frac{dg}{dx} \right|_{x=\mu_X} \right)^2 \sigma^2{}_X$$

Veamos la exactitud de la fórmula analítica.

$\mu_X = 1/2$

$\sigma^2{}_X = 1/12$

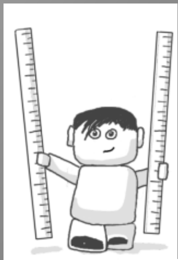
$$\mu_Y = \int_0^1 y\, 3\, y^2\, dy = 3/4$$

$$\sigma^2{}_Y = \mathbb{E}\{y^2\} - (\mu_Y)^2$$

$$\mathbb{E}\{y^2\} = \int_0^1 y^2\, 3\, y^2\, dy = 3/5$$

$$\sigma^2{}_Y = \frac{3}{5} - \left(\frac{3}{4}\right)^2 = 0.0375$$

$$\left(\left. \frac{dg}{dx} \right|_{x=\mu_X} \right)^2 \sigma^2{}_X = \frac{1}{3} \left(\frac{1}{2}\right)^{-4/3} \frac{1}{12} = 0.0233$$



Suceso de Bernoulli

A

$\longrightarrow 1$

B

$\longrightarrow 0$

Supongamos un suceso aleatorio con dos posibles resultados A y B. Asignaremos el valor 1 a la variable aleatoria para el evento A y el cero para B.

$p(A) = p$

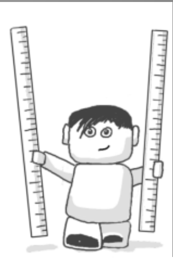
$p(B) = 1 - p = q$

Para este tipo de variable aleatoria tendremos que

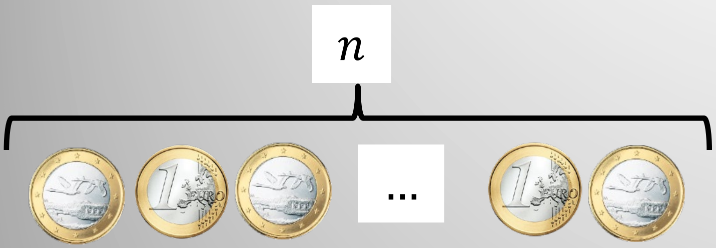
$$\mathbb{E}\{x\} = \sum_{i=1}^{2} p_i \, x_i = 1\, p + 0\, q = p$$

$$\sigma^2 = \mathbb{E}\{(x - \mathbb{E}\{x\})^2\} = \sum_{i=1}^{2} p_i \, (x_i - p)^2 = p\,(1-p)^2 + \; q\,(0-p)^2 = p(1-p) = pq$$

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA



Distribución binomial



Supongamos un suceso aleatorio formado por n eventos de Bernoulli estadísticamente independientes.

$$X(\Omega) = \{0,1,2,3, \dots, n\}$$

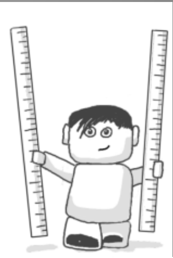
Asignamos como valor de la variable aleatoria el número de veces que obtenemos el suceso A.

$$\binom{n}{r} = \frac{n!}{r!\,(n-r)!}$$

$$p(x = r) = \binom{n}{r} p^r q^{n-r}$$

Con esta definición se cumple la condición de normalización:

$$\sum_{r=0}^{n} p(x = r) = \sum_{r=0}^{n} \binom{n}{r} p^r q^{n-r} = (p+q)^n = 1$$



UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Distribución binomial

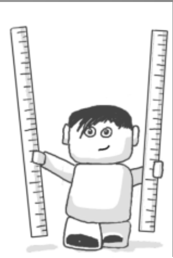
n

$X(\Omega) = \{0,1,2,3,\dots,n\}$

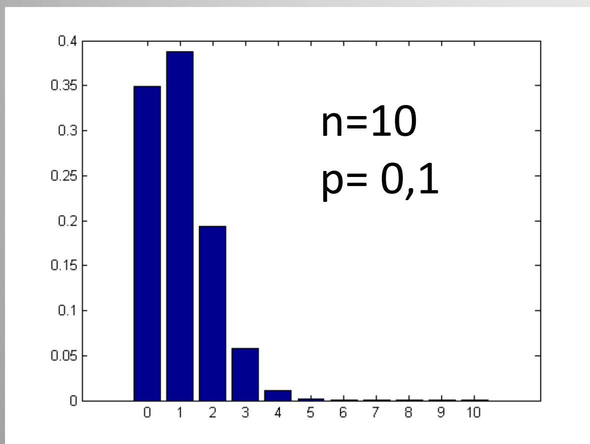
Podemos calcular el valor esperado y la varianza de una distribución binomial.

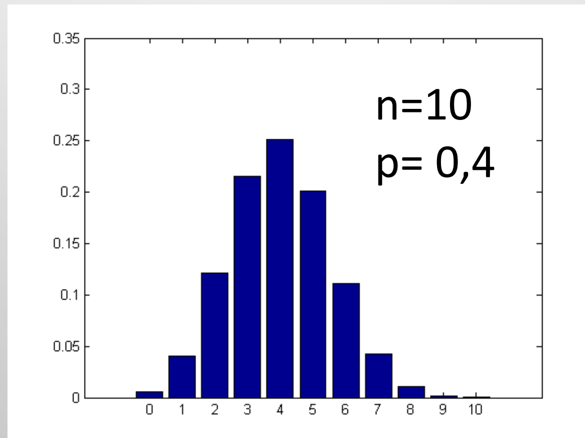
$$\mathbb{E}\{x\} = \sum_{r=0}^{n} r\, p(r) = \sum_{r=0}^{n} r \binom{n}{r} p^r q^{n-r} = np$$

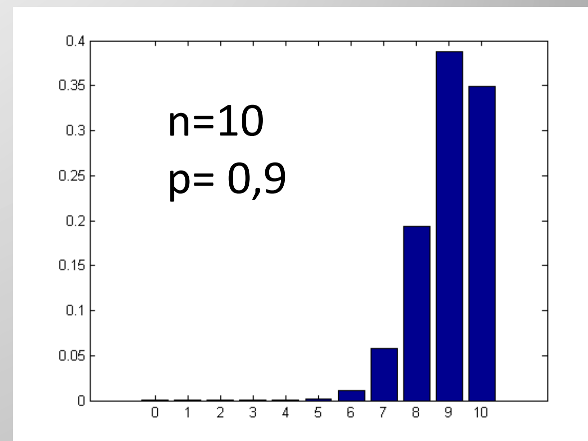
$$\sigma^2 = \sum_{r=0}^{n} r^2\, p(r) - (np)^2 = (np)^2 - np^2 + np - (np)^2 = npq$$

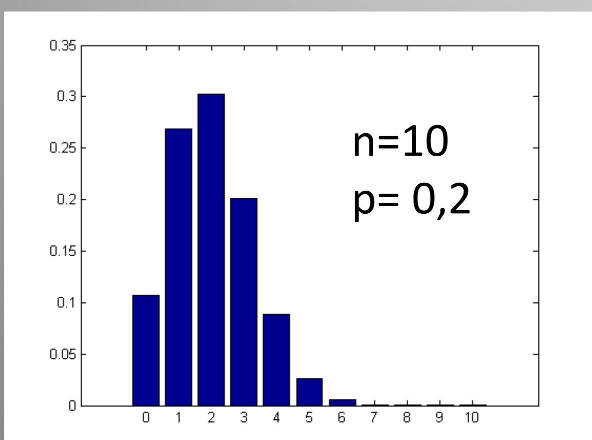


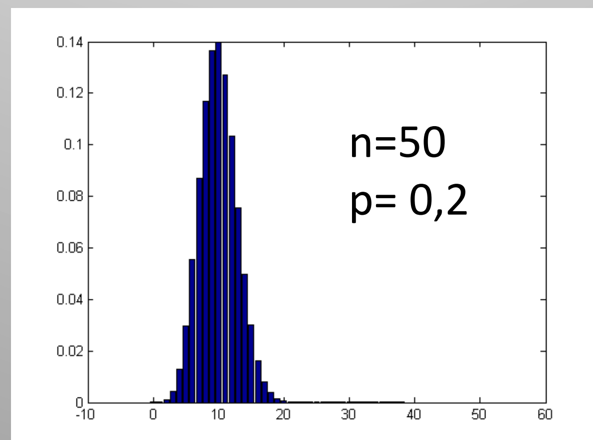
Distribución binomial

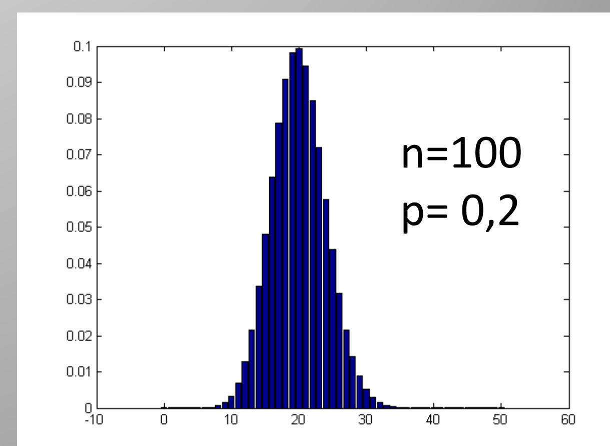




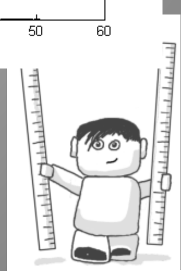








USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA



Distribución binomial

El sistema de aviónica de un fabricante de aviones de pasajeros presenta redundancia mediante lógica mayoritaria copiando el programa de control en tres memorias equivalentes. Si se considera que la probabilidad de fallo individual durante un vuelo de un módulo de memoria es de 10^{-6} ¿cuál es la probabilidad de que fallen simultáneamente más de uno de los tres módulos de control?

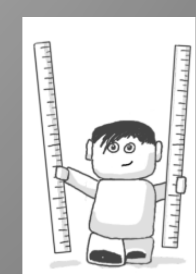
$$p = 10^{-6}$$

$$p(x>1) = 1 - p(0) - p(1) = 1 - \binom{3}{0} p^0 q^3 - \binom{3}{1} p^1 q^{3-1} = \binom{3}{2} p^2 q^1 + \binom{3}{3} p^3$$

$$p(x>1) = 1 - (1-p)^3 - 3\,p^1(1-p)^2 = 3\,p^2(1-p) + p^3 = 3p^2 - 2p^3$$

$$p(x>1) = 3 \times 10^{-12}$$

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA



Grandes números

Si realizamos un experimento aleatorio n veces para estimar la probabilidad de un suceso A, tendremos que la frecuencia observada es

$$f_A = \frac{x}{n}$$

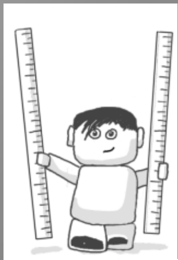
El valor esperado de esta frecuencia se corresponde con la probabilidad que queremos estimar ya que

$$\mathbb{E}\left\{\frac{x}{n}\right\} = \frac{1}{n}\mathbb{E}\{x\} = \frac{1}{n}\,np = p$$

$$\sigma^2(f_A) = \sigma^2\left(\frac{x}{n}\right) = \frac{1}{n^2}\,\sigma^2(x) = \frac{1}{n^2}npq = \frac{pq}{n}$$

Si no existen otras componentes de incertidumbre o variabilidad en la determinación de *x*, entonces se cumple que:

$$\sigma(f_A) \sim \frac{1}{\sqrt{n}} \xrightarrow[n\to\infty]{} 0$$



Grandes números

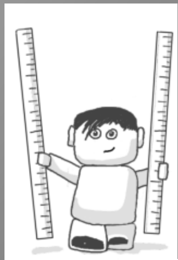
Se sabe que cierta enfermedad tiene una incidencia entorno al 1/200 de la población. Si queremos determinar esta probabilidad de incidencia con un 1% de incertidumbre (desviación típica) ¿Cuál es el tamaño de la muestra de población que debemos estudiar?

$$\sigma(f_A) = 0{,}01\, f_A = \frac{0{,}01}{200}$$

$$\sigma^2(f_A) = \frac{pq}{n} = \frac{1}{200}\left(1 - \frac{1}{200}\right)\frac{1}{n}$$

$$\frac{1}{200}\left(\frac{199}{200}\right)\frac{1}{n} = \frac{10^{-4}}{200^2}$$

$$n = \frac{199}{10^{-4}} = 1{,}99 \times 10^6$$



Distribución geométrica o de Pascal

Se trata de evaluar el número de veces que se produce el experimento aleatorio hasta que el suceso A con probabilidad *p* no ocurra.

1 **2** **…** **r-1** **r**

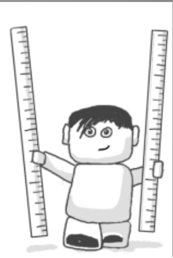
probabilidad de éxito hasta la tirada r

$$p(x=r)=\ p^{r-1}\,q$$

$$X(\Omega)=\{1,2,3,\dots\}$$

$$\sum_{r=1}^{\infty} p(x=r)=\sum_{r=1}^{\infty} p^{r-1}\,q=q\sum_{r=1}^{\infty} p^{r-1}=\ q\,\frac{1}{1-p}=1$$

Comprobamos que la probabilidad está correctamente definida usando la suma de la progresión geométrica.



Distribución geométrica o de Pascal

Podremos también evaluar el valor esperado de la variable de esta distribución así como su varianza.

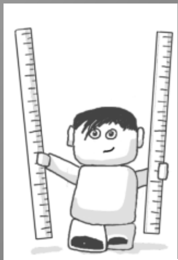
$$\mu = \mathbb{E}\{x\} = \sum_{r=1}^{\infty} r\, p(x=r) = \sum_{r=1}^{\infty} r\, p^{r-1}\,(1-p)$$

Teniendo en cuenta que la suma de probabilidades es igual a la unidad:

$$\sum_{r=1}^{\infty} p^{r-1}\,(1-p) = 1 = \sum_{s=0}^{\infty} p^{s}\,(1-p)$$

$$0 = \frac{d}{dp}\sum_{s=0}^{\infty} p^{s}\,(1-p) = \sum_{s=1}^{\infty} s\, p^{s-1}\,(1-p) - \sum_{s=0}^{\infty} p^{s}$$

$$\mu = \mathbb{E}\{x\} = \sum_{r=1}^{\infty} r\, p^{r-1}\,(1-p) = \frac{1}{q}$$



Distribución geométrica o de Pascal

La varianza pode estimarse mediante

$$\sigma^2 = \mathbb{E}\{x^2\} - (\mathbb{E}\{x\})^2 = \sum_{r=1}^{\infty} r^2 p^{r-1}(1-p) - \frac{1}{q^2}$$

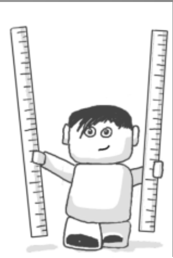
Podemos reescribir esta suma como:

$$\sum_{r=1}^{\infty} r^2 p^{r-1}(1-p) = \sum_{r=1}^{\infty} r(r-1)p^{r-1}(1-p) + \underbrace{\sum_{r=1}^{\infty} r\,p^{r-1}(1-p)}_{\frac{1}{q}}$$

$$\sum_{r=1}^{\infty} r\,p^{r-1}(1-p) = \frac{1}{1-p} \quad \Rightarrow \quad \sum_{r=1}^{\infty} r(r-1)p^{r-1}(1-p) = \frac{2p}{(1-p)^2}$$

De donde $\sigma^2 = \dfrac{p}{q^2}$

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA



Distribución geométrica o de Pascal

Supongamos que jugamos a la ruleta rusa con una pistola de revólver con 6 alojamientos y una sola bala. ¿Cuál es la probabilidad de que sobrevivamos a 1, 2, 3 y 4 disparos si giramos el tambor o no lo hacemos?

probabilidad de éxito hasta la tirada r $p(x = r) = p^{r-1}\, q$

La probabilidad de sobrevivir a k intentos girando el tambor

$$p(x > k) = 1 - q - p\,q - \cdots - p^{k-1}\,q \qquad p = 5/6$$

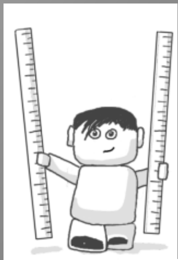
$p(x > 1) = 0{,}83$ $p(x > 2) = 0{,}69$ $p(x > 3) = 0{,}58$ $p(x > 4) = 0{,}48$ $p(x > 5) = 0{,}40$

La probabilidad de sobrevivir a si no giramos el tambor

$$p(x > k) = \frac{5}{6}\,\frac{4}{5}\,\frac{3}{4}\,\dots\frac{6-k}{6-k+1}$$

$p(x > 1) = 0{,}83$ $p(x > 2) = 0{,}67$ $p(x > 3) = 0{,}50$ $p(x > 4) = 0{,}33$ $p(x > 5) = 0{,}17$

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA



Distribución de Poisson

Si consideramos una distribución de probabilidad binomial y realizamos el límite de esta distribución con $\boldsymbol{n \to \infty}$ manteniendo el valor esperado constante $\mathbb{E}\{\boldsymbol{x}\} = \boldsymbol{p\,n} = \boldsymbol{\lambda} = \boldsymbol{cte}$

$$p(x=r) = \binom{n}{r} p^r q^{n-r} = \binom{n}{r} \left(\frac{\lambda}{n}\right)^r \left(1 - \frac{\lambda}{n}\right)^{n-r}$$

Podremos escribir:

$$\lim_{\substack{n\to\infty \\ p\to 0 \\ np=\lambda=cte}} p(x=r) = \frac{\lambda^r}{r!} \lim_{n\to\infty} \left\{ \frac{n(n-1)\ldots(n-r+1)}{\left(1-\frac{\lambda}{n}\right)^r n^r} \left(1-\frac{\lambda}{n}\right)^n \right\} =$$

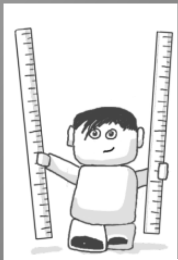
$$= \frac{\lambda^r}{r!} \lim_{n\to\infty} \left\{ \frac{n(n-1)\ldots(n-r+1)}{(n-\lambda)^r} \left(1-\frac{\lambda}{n}\right)^n \right\}$$

Por otra parte:

$$\lim_{n\to\infty} \left\{ \left(\frac{n}{n-\lambda}\right) \left(\frac{n-1}{n-\lambda}\right) \ldots \left(\frac{n-r+1}{n-\lambda}\right) \right\} = 1$$

$$\lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda}$$

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA



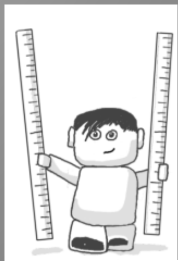
Distribución de Poisson

Mediante este límite obtenemos la expresión de la probabilidad de Poisson,

$$p(x=r) = \lim_{\substack{n\to\infty \\ p\to 0 \\ np=\lambda=cte}} \binom{n}{r}\left(\frac{\lambda}{n}\right)^{r}\left(1-\frac{\lambda}{n}\right)^{n-r} = \frac{\lambda^{r}}{r!}e^{-\lambda}$$

$$r = 0,1,2,\ldots$$

Esta distribución de probabilidad uniparamétrica corresponde a la realización de muchos experimentos aleatorios donde la probabilidad individual del suceso de interés es muy pequeña siendo el valor esperado de sucesos detectable.



Distribución de Poisson

A partir de la expresión de la probabilidad de Poisson, que es una distribución discreta, podremos establecer su correcta normalización y su valor esperado y varianza.

$$p(x = r) = \frac{\lambda^r}{r!} e^{-\lambda}$$

$$r = 0,1,2, \dots$$

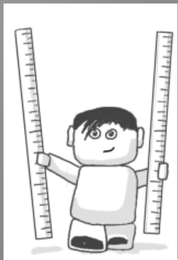
$$\sum_{r=0}^{+\infty} p(x = r) = \sum_{r=0}^{+\infty} \frac{\lambda^r}{r!} e^{-\lambda} = e^{-\lambda} \sum_{r=0}^{+\infty} \frac{\lambda^r}{r!} = e^{-\lambda} e^{\lambda} = 1$$

Normalización

$$\mathbb{E}\{x\} = \sum_{r=0}^{+\infty} r\, p(x = r) = \sum_{r=0}^{+\infty} r\, \frac{\lambda^r}{r!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{r=1}^{+\infty} \frac{\lambda^{r-1}}{(r-1)!} = \lambda\, e^{-\lambda} e^{\lambda} = \lambda$$

El valor esperado de la variable es $\boldsymbol{\lambda}$

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA



Distribución de Poisson

El cálculo de la varianza de la distribución:

$$\sigma^2(x) = \mathbb{E}\{x^2\} - (\mathbb{E}\{x\})^2 = \sum_{r=0}^{+\infty} r^2\, p(x=r) - \lambda^2 = \sum_{r=0}^{+\infty} r^2\, \frac{\lambda^r}{r!} e^{-\lambda} - \lambda^2$$

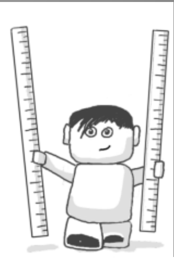
$$\sum_{r=0}^{+\infty} r^2\, \frac{\lambda^r}{r!} e^{-\lambda} = e^{-\lambda} \sum_{r=0}^{+\infty} [r(r-1)+r]\, \frac{\lambda^r}{r!} = e^{-\lambda}\left[\lambda^2 \sum_{r=2}^{+\infty} \frac{\lambda^{r-2}}{(r-2)!} + \lambda \sum_{r=1}^{+\infty} \frac{\lambda^{r-1}}{(r-1)!}\right]$$

$$\sum_{r=0}^{+\infty} r^2\, \frac{\lambda^r}{r!} e^{-\lambda} = e^{-\lambda}\left[\lambda^2\, e^{\lambda} + \lambda\, e^{\lambda}\right] = \lambda^2 + \lambda$$

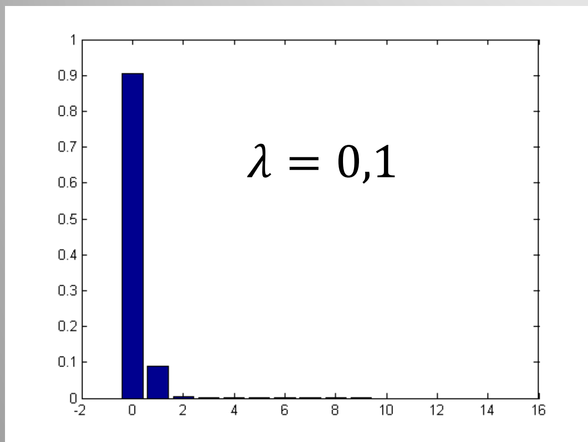
$$\sigma^2(x) = \lambda^2 + \lambda - \lambda^2$$

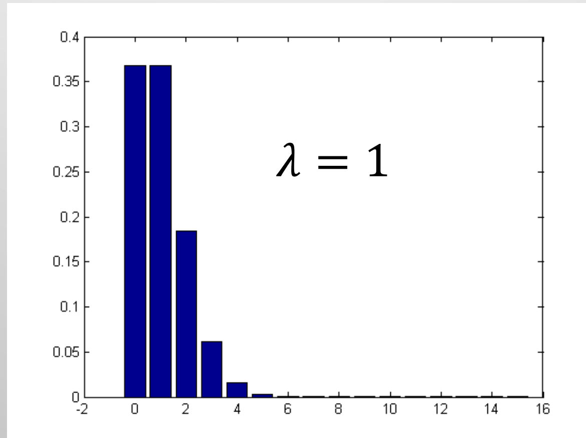
En el caso de la distribución de Poisson (uniparamétrica) se cumple

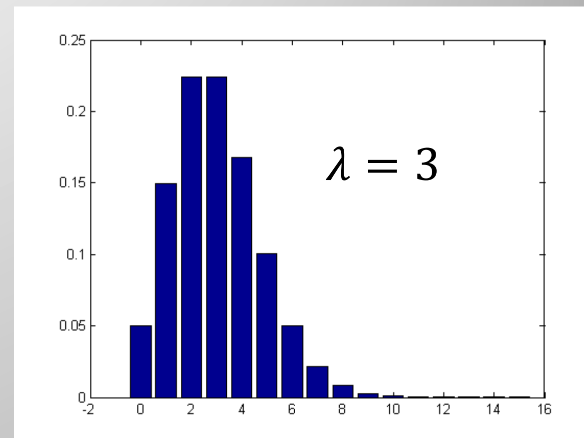
$$\mu = \sigma^2 = \lambda$$

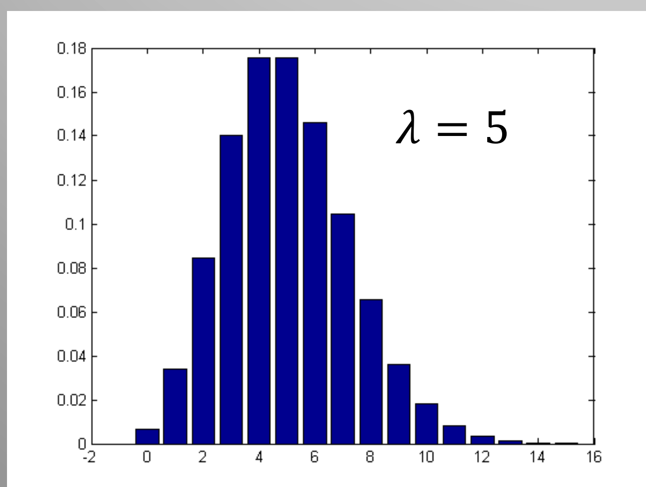


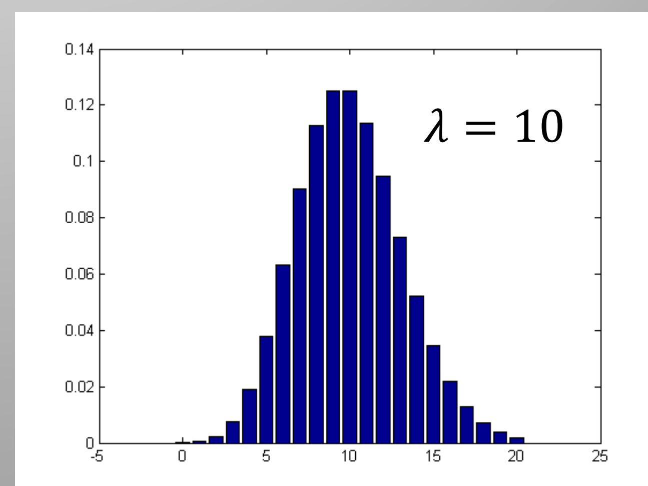
Distribución de Poisson

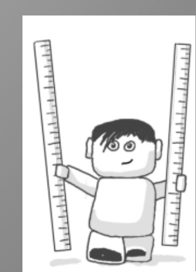












Distribución de Poisson

El número medio de accidentes de tráfico por hora en una cierta ciudad es de 1/20. Calcula cúal es la probabilidad de que:

a) No se produzca ningún accidente en una hora

b) La probabilidad de que se produzcan más de dos accidentes en una hora

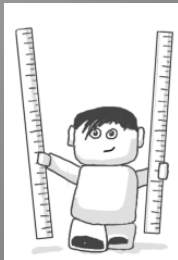
$$p(x=0) = e^{-\lambda} = 0.9512$$

$$p(x>2) = 1 - p(x=0) - p(x=1) - p(x=2)$$

$$p(x>2) = 1 - e^{-\lambda} - \lambda\, e^{-\lambda} - \frac{\lambda^2}{2} e^{-\lambda} = 1 - \left(1 + \lambda + \frac{\lambda^2}{2}\right) e^{-\lambda}$$

$$p(x>2) = 2 \times 10^{-5}$$

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA



Binomial-Poisson-Gauss

Binomial

(n,p)

$\mu = np$

$\sigma^2 = npq$

Discreta

$n \to \infty$

$np = \lambda = cte$

$p \ll 0{,}1$

Poisson

(λ)

$\mu = \lambda$

$\sigma^2 = \lambda$

Discreta

$n \to \infty$

$\lambda \gg 5$

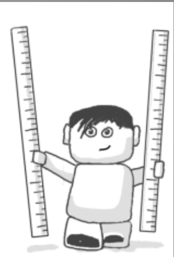
$\lambda \to \infty$

Normal/Gauss

(μ, σ)

Continua

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA



Distribución normal o gaussiana

Una variable aleatoria continua se dice que tiene una distribución de probabilidad gaussiana cuando su densidad de probabilidad se puede escribir como:

$$f(x) = \frac{1}{\sqrt{2\pi}\ \sigma}\ e^{-\frac{(x-\mu)^2}{2\,\sigma^2}} \quad ; \quad \sigma > 0 \quad x \in (-\infty, +\infty)$$

Y su función de distribución de probabilidad:

$$F(x) = \frac{1}{\sqrt{2\pi}\ \sigma} \int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\,\sigma^2}}\, dt; \;\; \sigma > 0 \quad x \in (-\infty, +\infty)$$

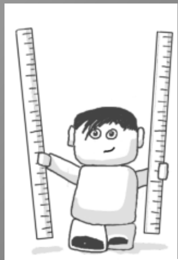
Los parámetros de la distribución σ, μ cumplen que:

$$\mathbb{E}\{x\} = \ \mu \qquad \mathbb{E}\{(x - \mathbb{E}\{x\})^2\} = \ \sigma^2$$

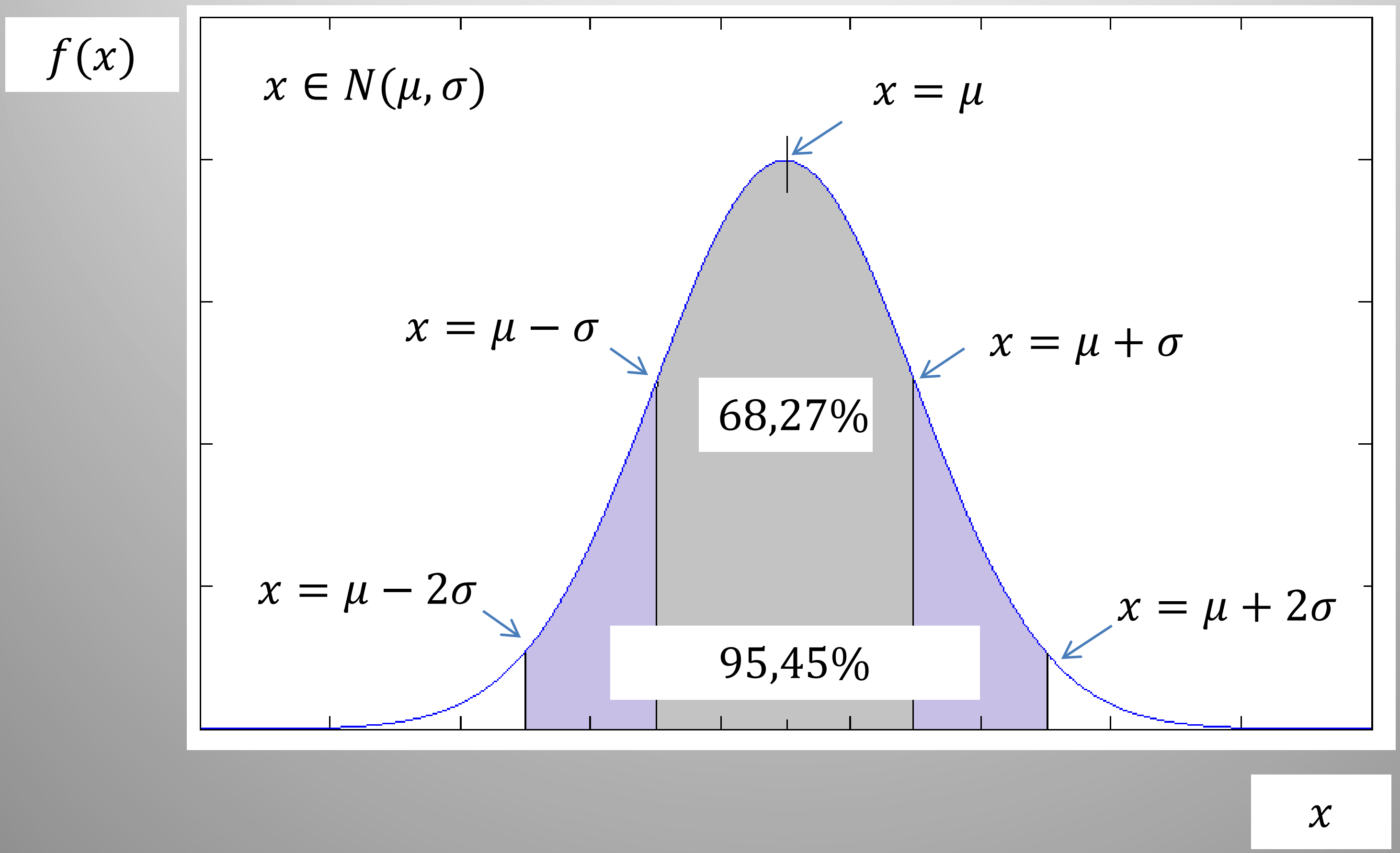
Es habitual denotar una variable normal o gaussiana como

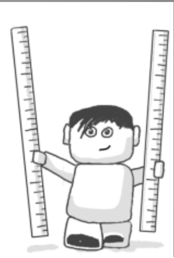
$$x \in N(\mu, \sigma)$$

Es una distribución de probabilidad simétrica respecto al punto x=μ, con máximo y moda en el mismo punto.



Distribución normal o gaussiana

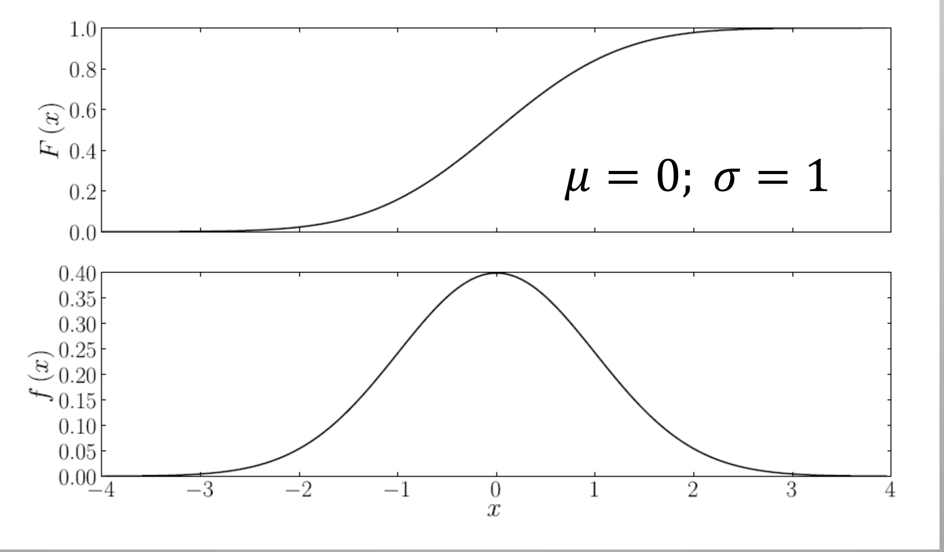




Distribución normal o gaussiana estándar

En el caso particular de que μ= 0 y σ=1, se denomina distribución de probabilidad normal estándar

$$f(z) = \frac{1}{\sqrt{2\pi}}\ e^{-\frac{z^2}{2}} \quad ; \quad z \in (-\infty, +\infty)$$



Es habitual denotar una variable normal estándar como

$$z \in N(0,1)$$

Función generatriz de momentos para la normal estándar

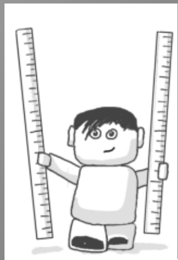
Considerando la distribución normal estándar, tendremos que

$$f(z) = \frac{1}{\sqrt{2\pi}} \; e^{-\frac{z^2}{2}} \quad ; \quad z \in (-\infty, +\infty)$$

$$\mathcal{M}_X(t) = \mathbb{E}\{e^{tX}\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{z^2}{2}} \; e^{zt} \, dz$$
$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\left[-\frac{1}{2}(z^2 - 2zt + t^2)\right] \exp\left[\frac{t^2}{2}\right] dz$$

$$\mathcal{M}_X(t) = \exp\left[\frac{t^2}{2}\right]$$

De donde se pueden obtener los momentos de orden arbitrario de la distribución normal estándar mediante la derivación.



Distribución normal

Una fábrica de rodamientos de acero produce bolas con un diámetro que puede describirse bien mediante una distribución normal con media 10 mm y desviación típica 0,2 mm. Si se considera que el diámetro aceptable de las bolas de rodamientos es de 9,7 mm a 10,1 mm , calcula el porcentaje de bolas rechazadas.

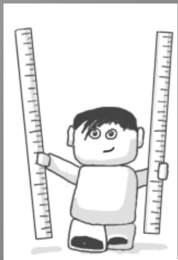
$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{9{,}7 - 10}{0{,}2} = -1{,}5$$

$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{10{,}1 - 10}{0{,}2} = 0{,}5$$

$$p(z < z_1) = p(z < -1{,}5) = 0{,}067$$

$$p(z > z_2) = p(z > 0{,}5) = 0{,}309$$

$$p(z < z_1 \text{ ó } z_2 < z) = 0{,}067 + 0{,}309 = 0{,}376$$



Curtosis de la distribución normal

Utilizando la función generatriz de momentos es relativamente fácil evaluar la curtosis de la distribución normal:

$$g = \frac{M_4(\mu)}{\sigma^4}$$

En este caso particular:

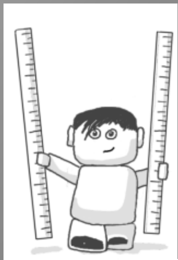
$$M_r(0) = \left. \frac{d^r exp\left[\frac{t^2}{2}\right]}{dt^r} \right|_{t=0}$$

$$\frac{d^4 exp\left[\frac{t^2}{2}\right]}{dt^4} = 3\ exp\left[\frac{t^2}{2}\right] + 6t^2 exp\left[\frac{t^2}{2}\right] + t^4 exp\left[\frac{t^2}{2}\right]$$

Por lo cual podemos deducir que se cumple:

$$g = 3$$

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA



Bibliografía:

Capítulos 4 y 5. “Probabilidad y estadística” George C. Canavos, Ed. Mc Graw-Hill

Capítulo 5. “Fundamentos de estadística” Daniel Peña, Alianza Editorial

Capítulo 2. “Tratamiento de datos físicos” Faustino Gómez, Luis M Varela, USC

Capítulo 5. “Statistical and Computational Methods for Scientists and Engineers” Siegmund Brandt, Springer

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA