Las Malvinas
son argentinas

ESTADÍSTICA
DIAPOSITIVAS CON EJEMPLOS

Unidad 4 – *Modelos Probabilísticos*

Ingeniería en Informática

Año 2022

Prof. Juan Pablo Taulamet

consultas: taulamet@unl.edu.ar

Apache

Un Sitio Web es atendido por cuatro servidores para distribuir la carga. Cuando llega un pedido sólo uno se encarga de responder a la solicitud y los otros funcionan como respaldos por si falla el servidor activo. La probabilidad de que una solicitud al sitio Web genere un fallo en el servidor activo es 0,01. Si cada solicitud es una prueba independiente; ¿Cuál será el número medio de fallos?

Ensayo de Bernoulli

Éxito o Fracaso

Llega una solicitud: “Experimento”

- A: Se responde correctamente
- ~A: Se produce un fallo

Datos

Un Sitio Web es atendido por cuatro servidores … la probabilidad de que una solicitud al sitio Web genere un fallo en el servidor activo es 0,01.

Paradójicamente: Éxito → Fallo

- 4 Servidores
- P(E) = 0,01

Interrogante

¿Cuál será el número medio de fallos?

Si X: Número de Fallos

Entonces E(X): Número medio de Fallos

Planteo

X: Número de Fallos

X $\rightarrow$ V.A Discreta

X $\rightarrow$ Unidimensional

X ~ ?

Planteo

X: Número de Fallos (V.A. Discreta)

X → V.A Discreta

X → Unidimensional

X ~ Binomial(n=4, p=0,01)

=> E(X) = n*p = 0,04

Rta: El número esperado de fallos es 0.

Un reciente estudio revela que el 40% de los adultos están a favor de un control de precios y salarios. Si se seleccionaran 5 adultos aleatoriamente:

a) Determine la probabilidad de que ninguno esté a favor del citado control.
b) Cuál es la probabilidad de que como máximo 3 estén a favor del control?
c) Cuál el valor esperado de adultos seleccionados que están a favor del control?

Ensayo de Bernoulli

Éxito o Fracaso

Experimento: Se le consulta a un adulto su opinión sobre el control de precios:

- A: El adulto está a favor
- ~A: El adulto está en contra

Datos

“...el 40% de los adultos están a favor de un control de precios y salarios.”

En este caso: Éxito → A Favor

E: El adulto está a favor del control

- 5 Adultos
- $P(E) = 0,40$

Interrogante a)

Si se seleccionan 5 adultos al azar:

¿Determine la probabilidad de que **ninguno** esté a favor del citado control?

Si X: “Número de adultos a favor”

$$P(X=0) = ?$$

Planteo a)

X: Número de adultos a favor

X → V.A Discreta

X → Unidimensional

X ~ ?

Planteo a)

X: Número de de adultos a favor (V.A. Discreta)

X → V.A Discreta

X → Unidimensional

X ~ Binomial(n=5, p=0,40)

$$P(X=0)=\binom{n}{x}\mathrm{p}^{x}(1-p)^{n-x}=\binom{5}{0}0,40^{0}*0,60^{5}$$

$$P(X=0)=0,078 \quad =binomdist(0;0,40;5;1)$$

Interrogante b)

Si se seleccionan 5 adultos al azar:

¿Determine la probabilidad de que **como máximo 3** estén a favor del citado control?

Si X: “Número de adultos a favor”

$$P(X \leq 3) = ?$$

Planteo b)

X: Número de de adultos a favor (V.A. Discreta)

X → V.A Discreta

X → Unidimensional

X ~ Binomial(n=5, p=0,40)

$$P(X \leq 3) = \sum_{i=0}^{3} \binom{n}{x_i} \; \mathrm{p}_i^x (1-p)^{n-x_i}$$

$$P(X \leq 3) = 0,91 \qquad = binomdist(3; 0,40; 5; 1)$$

Interrogante c)

Si se seleccionan 5 adultos al azar:

¿Cuál es el **valor esperado** de adultos a favor del control?

Si X: “Número de adultos a favor”

$$E(X) = ?$$

Planteo c)

Si se seleccionan 5 adultos al azar:

¿Cuál es el **valor esperado** de adultos a favor del control?

Si X: “Número de adultos a favor”

Si X ~Binom

E(X) = n*p = 5*0,40 = 2

Respuestas

a) La probabilidad de que ninguno esté a favor del citado control es 0,078.

b) La probabilidad de que como máximo 3 estén a favor del control es 0,91.

c) El valor esperado de adultos seleccionados que están a favor del control es 2.

Problema

El número de días entre la facturación y el pago de cuentas corrientes en un negocio tiene una distribución con los siguientes parámetros: media = 18 días y desvío = 3 días.

¿ Qué proporción de las facturas será pagada

a) entre 12 y 18 días?

b) en menos de 8 días?

c) ¿Dentro de cuántos días estará pagado el 99.5% de las facturas?

Datos

X: “Tiempo en días entre la facturación y el pago”

- $X \rightarrow$ V. A. Continua
- $X \rightarrow$ Unidimensional
- $E(X) = 18$
- $D(X) = 3$
- $X \sim ?$

Interrogantes

¿ Qué proporción de las facturas será pagada

- entre 12 y 18 días?
- en menos de 8 días?
- ¿Dentro de cuántos días estará pagado el 99.5% de las facturas?

Interrogantes

¿ Qué proporción de las facturas será pagada...

- **entre 12 y 18 días?**
- en menos de 8 días?
- ¿Dentro de cuántos días estará pagado el 99.5% de las facturas?

Interrogantes

¿ Qué proporción de las facturas será pagada

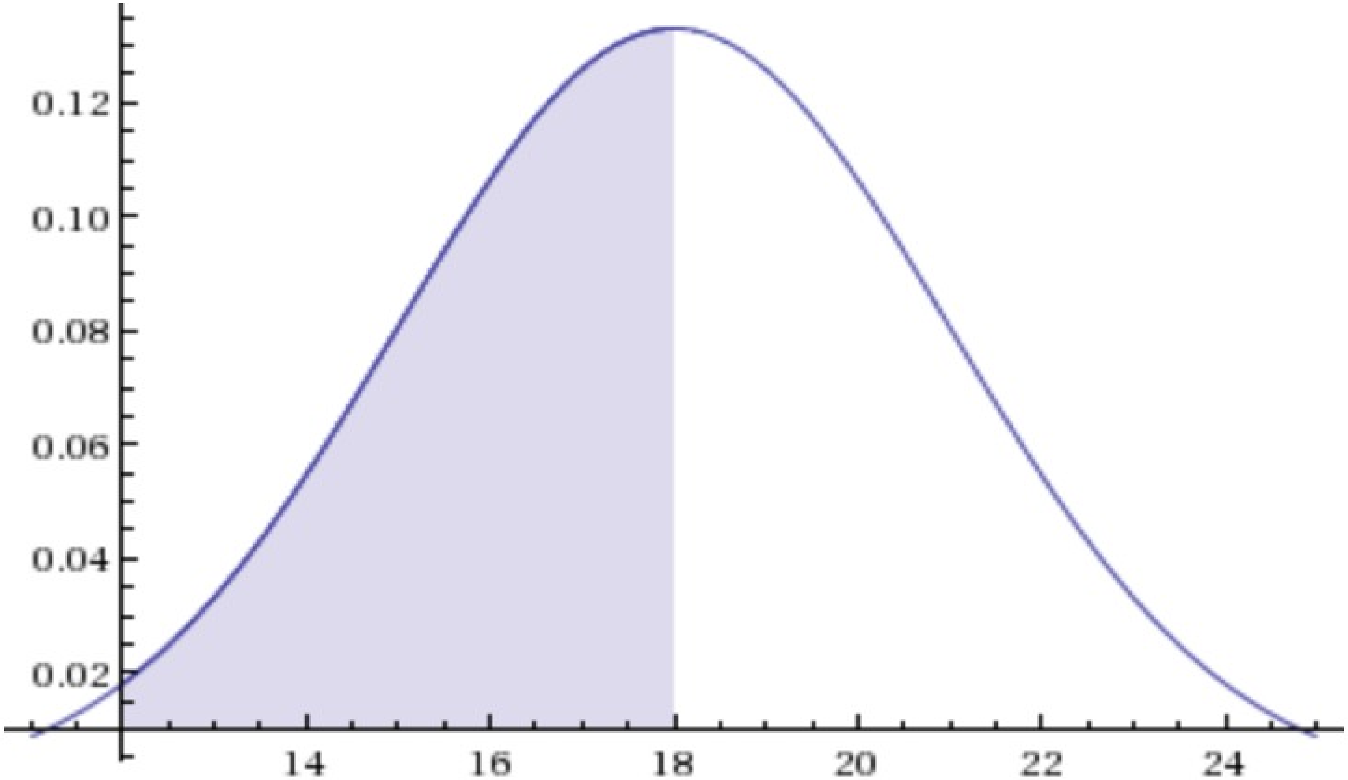
- **entre 12 y 18 días?**

$$P\left(12 \leq X \leq 18\right) = ?$$

Planteo a)

¿ Qué proporción de las facturas será pagada entre 12 y 18 días?

$$\mathrm{P}(12 \leq X \leq 18) = P(X \leq 18) - P(X \leq 12) = F(18) - F(12)$$



Planteo a)

¿ Qué proporción de las facturas será pagada entre 12 y 18 días?

$\mathrm{P}(12 \leq X \leq 18) = P(X \leq 18) - P(X \leq 12) =$
$F(18) - F(12) =$

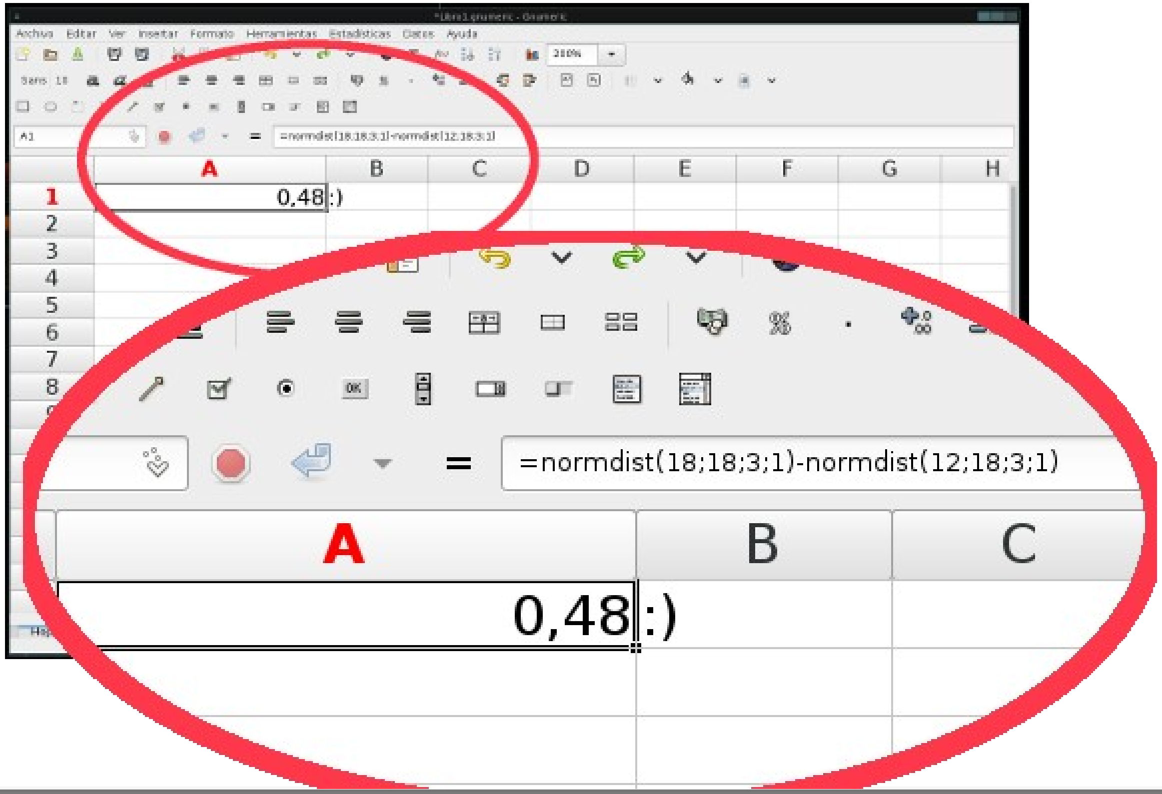
en Gnumeric
normdist(18; 18; 3; 1) - normdist(12; 18; 3; 1)

siendo
$\mathrm{P}(\mathrm{X} \leq x) = F(x) = normdist(x; E(X); D(X); 1)$

Planteo a)

¿ Qué proporción de las facturas será pagada entre 12 y 18 días?

$$\mathrm{P}(12 \leq X \leq 18) = P(X \leq 18) - P(X \leq 12) =$$
$$= normdist(18; 18; 3; 1) - = normdist(12; 18; 3; 1)$$



Análogamente b)

¿ Qué proporción de las facturas será pagada

- entre 12 y 18 días?
- **en menos de 8 días?**
- ¿Dentro de cuántos días estará pagado el 99.5% de las facturas?

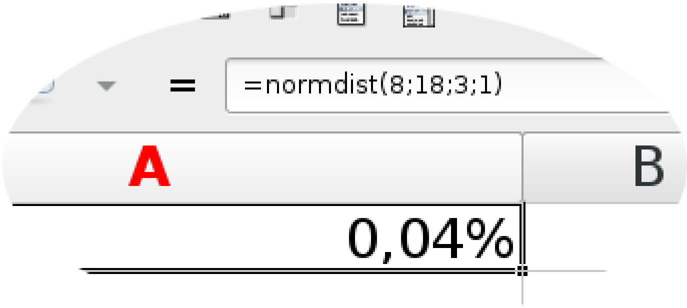
Planteo b)

¿ Qué proporción de las facturas será pagada en menos de 8 días?

$$\mathrm{P}(\mathrm{X} \leq 8) = P(X \leq 8) = F(8)$$

en Gnumeric

normdist(8; 18; 3; 1)



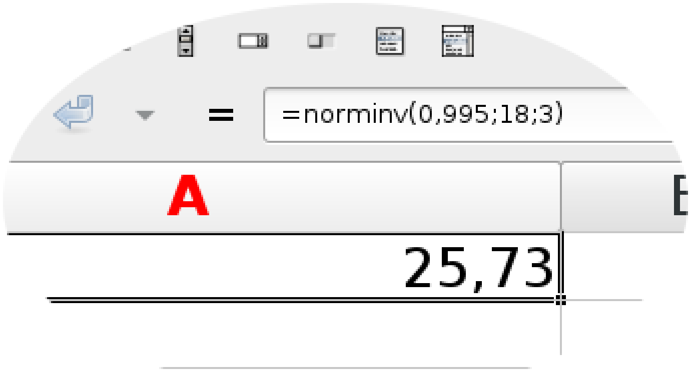
Planteo c)

- **¿Dentro de cuántos días estará pagado el 99.5% de las facturas?**

$$\mathrm{P}(\mathrm{X} \leq x) = F(X = x) = 0,995; x = ?$$

en Gnumeric

norminv(0.995; 18; 3)



Respuestas

¿ Qué proporción de las facturas será pagada

- entre 12 y 18 días? → 48%
- en menos de 8 días? → 0,04%
- ¿Dentro de cuántos días estará pagado el 99.5% de las facturas? → 25 días y 18 hs. aproximadamente.

Problema

Un camión transporta novillos cuyos pesos tienen una distribución con media 490 kg y desvío estándar 20 kg. Por disposiciones reglamentarias no puede llevar más de 10.000 kg de carga. Determine Ud el número máximo de novillos que puede llevar a efectos de que la carga máxima se supere sólo con un 5% de probabilidad.

Problema

Un camión transporta ***Ingenieros*** cuyos pesos tienen una distribución con media 490 kg y desvío estándar 20 kg. Por disposiciones reglamentarias no puede llevar más de 10000 kg de carga. Determine Ud el número máximo de novillos que puede llevar a efectos de que la carga máxima se supere sólo con un 5% de probabilidad.

Problema

Un camión transporta Novillos ***Expertos en Informática*** cuyos pesos tienen una distribución con media 490 kg y desvío estándar 20 kg. Por disposiciones reglamentarias no puede llevar más de 10.000 kg de carga.

Determine el número máximo de novillos que puede llevar a efectos de que la carga máxima se supere sólo con un 5% de probabilidad.

Datos

X: “Peso en Kg. de un Novillo Experto”

- X $\rightarrow$ V. A. Continua
- X $\rightarrow$ Unidimensional
- $E(X) = 490$
- $D(X) = 20$
- $X \sim ?$

Interrogante

Número máximo de novillos que puede llevar a efectos de que la carga máxima se supere sólo con un 5% de probabilidad.

$$n = ?$$

Datos

X: “Peso en Kg. de un Novillo Experto”

Y: “Peso en Kg. de **una manada de n** Novillos”

$$\mathrm{E(X)} = 490$$
$$\mathrm{E(Y)} = \mathrm{n}{*}490$$
$$\mathrm{D(X)} = 20$$
$$\mathrm{V(Y)} = \mathrm{n}^2.V(X)$$
$$D(Y) = \sqrt{n^2.V(X)} = n.D(X)$$

Planteo

X: “Peso en Kg. de un Novillo”

Y: “Peso en Kg. de una manada n de Novillos”

$$\mathrm{X} \sim N(\mu = 490, \sigma = 20)$$
$$Y \sim N(\mu = n * 490, \sigma = n * 20)$$

Resolución

Número máximo de novillos que puede llevar a efectos de que la carga máxima se supere sólo con un 5% de probabilidad.

Y: “Peso en Kg. de una manada de Novillos”

$$\mathrm{P}(\mathrm{Y} \geq 10.000) = 5\% \Rightarrow P(Y \leq 10.000) = 95\%$$

$$\text{Estandarizando: } \; z = \frac{x-\mu}{\sigma}$$

$$P(\frac{Y-\mu_Y}{\sigma_Y} \leq \frac{10.000-n.490}{n.20}) = 95\%$$

$$P(z \leq \frac{500}{n} - 24,5) = 95\%$$

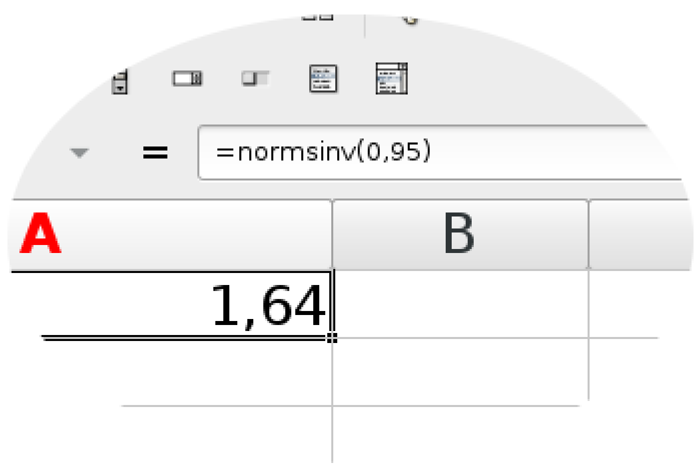
Obteniendo Z

$$\mathrm{P(Y \leq 10.000) = 95\%}$$

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

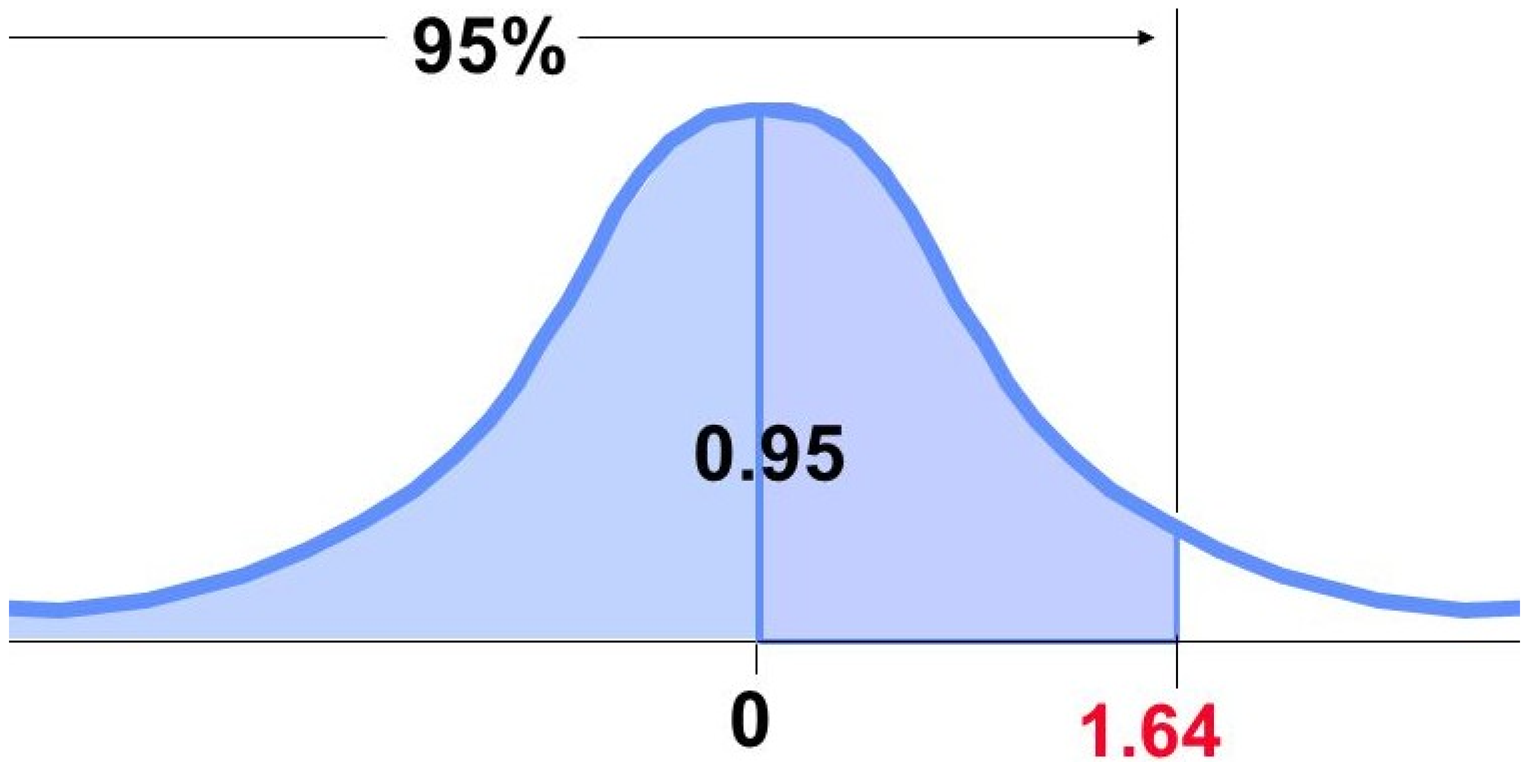
$$P\left(\frac{Y-\mu_Y}{\sigma_Y} \leq \frac{10.000-n.490}{n.20}\right) = 95\% = P\left(z \leq \frac{500}{n} - 24,5\right)$$

$$\Rightarrow \frac{500}{n} - 24,5 = z_{95\%} = normsinv(0,95)$$



Despejamos n

$$P\left(\frac{Y-\mu}{\sigma} \leq \frac{10.000-n.490}{n.20}\right) = 95\% = P\left(z \leq \frac{500}{n} - 24,5\right)$$
$$\Rightarrow \frac{500}{n} - 24,5 = z_{95\%} = 1.64 \Rightarrow n \approx 19$$



Respuesta

El Número máximo de novillos que puede llevar a efectos de que la carga máxima se supere sólo con un 5% de probabilidad es 19.

Fiabilidad de sistemas

Una gran mayoría de los equipos no tienen una tasa de fallos constante: es más probable que fallen a medida que envejecen. En este caso la tasa de fallos es creciente. Aunque también es posible encontrar equipos con tasas de fallos decrecientes.

El término fiabilidad es descripto en el diccionario de la RAE como "probabilidad de buen funcionamiento de algo".

Fiabilidad de sistemas

El término fiabilidad es descripto en el diccionario de la RAE como "probabilidad de buen funcionamiento de algo".

Si representamos la fiabilidad con la función R(t) referida al tiempo *t,* la podremos definir utilizando la noción de función de probabilidad acumulada F, de la siguiente forma:

$$\mathrm{F(t) = P(T} \leq t) => R(t) = P(T > t)$$
$$\mathrm{R(t) = 1\ \text{-}F(t)}$$

Problema

Suponga que el tiempo de fallo en horas, de dos equipos, siguen distribuciones de Weibull con parámetros para el primero y segundo respectivamente:

$$\mathrm{k} = 1,\ \mu_0 = 20$$

y

$$\mathrm{k} = 0.5,\ \mu_0 = 10$$

Determinar:

a) **¿Qué equipo posee mayor fiabilidad sabiendo que ambos han sido usados durante 10 horas?**

b) ¿En qué momento debería producirse un fallo, para que la fiabilidad sea equivalente en ambos sistemas?

Distribución de Weibull

Recordando la función acumulativa de Weibull:

$$F(t) = 1 - e^{-\left(\frac{t}{\mu_0}\right)^k}$$

Tenemos la función de Fiabilidad:

$$R(t) = 1 - F(t) = e^{-\left(\frac{t}{\mu_0}\right)^k}$$

Calculando...

$P(T > 10)$ en ambos sistemas:

$$\mathrm{R(t)} = \mathrm{e}^{\,-\left(\frac{t}{\mu_0}\right)^k}$$

1 2

$$\mathrm{k} = 1$$
$$\mu_0 = 20$$

$$\mathrm{k} = 0.5$$
$$\mu_0 = 10$$

$$\mathrm{R(t)} = \mathrm{e}^{\,-\left(\frac{10}{20}\right)^1} =$$
$$= e^{-\frac{1}{2}} \approx 61\%$$

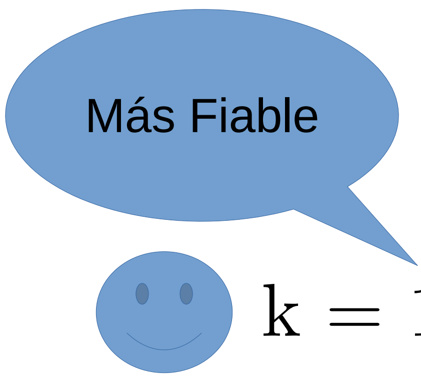
$$\mathrm{R(t)} = \mathrm{e}^{\,-\left(\frac{10}{10}\right)^{0.5}} =$$
$$= e^{-1} \approx 37\%$$

Calculando...

$$P(T > 10) \text{ en ambos sistemas:}$$

$$\mathrm{R(t)} = \mathrm{e}^{-\left(\frac{t}{\mu_0}\right)^k}$$

1 2



$$\mathrm{k} = 1 \qquad\qquad \mathrm{k} = 0.5$$

$$\mu_0 = 20 \qquad\qquad \mu_0 = 10$$

$$\mathrm{R(t)} = \mathrm{e}^{-\left(\frac{10}{20}\right)^1} = \\ = e^{-\frac{1}{2}} \approx 61\%$$

$$\mathrm{R(t)} = \mathrm{e}^{-\left(\frac{10}{10}\right)^{0.5}} = \\ = e^{-1} \approx 37\%$$

Segunda Parte

¿En qué momento debería producirse un fallo, para que la fiabilidad sea equivalente en ambos sistemas?

Segunda Parte

¿En qué momento debería producirse un fallo, para que la fiabilidad sea equivalente en ambos sistemas?

$$e^{-\left(\frac{t}{20}\right)^{1}} = e^{-\left(\frac{t}{10}\right)^{0.5}}\ ;t=?$$

$$ln(e^{-(\frac{t}{20})^{1}}) = ln(e^{-(\frac{t}{10})^{0.5}}) => (\frac{t}{20})^{1} = (\frac{t}{10})^{0.5}$$

$$=> t^{\frac{1}{2}} = \frac{20}{\sqrt{10}} => \quad t = 40$$

Verificación:

$$= 1 - weibull(40;1;20;1) = 13,53\%$$

$$= 1 - weibull(40;0,5;10;1) = 13,53\%$$

Respuestas

a) ¿Qué equipo posee mayor fiabilidad sabiendo que ambos han sido usados durante 10 horas?
Posee mayor fiabilidad el primer equipo.

b) ¿En qué momento debería producirse un fallo, para que la fiabilidad sea equivalente en ambos sistemas?
A las 40 horas.