HOJA 2

1. $$T = \frac{(O_1 - E_1)^2}{E_1} + \frac{(O_2 - E_2)^2}{E_2} \xrightarrow{d} \chi^2_1 \quad \text{si } n \to \infty$$

X v.a $\xrightarrow{\text{produce}}$ $\begin{cases} \alpha_1, \alpha_2 & \text{valores posibles} \\ p, 1-p & \text{probabilidades} \end{cases}$

Hay que ver que $T = Z^2$, donde $Z \sim N(0,1)$

X es similar a producir caras y cruces $\begin{cases} \nearrow \alpha_1 = \text{cara} = 1 \\ \searrow \alpha_2 = \text{cruz} = 0 \end{cases} \Rightarrow$

$\Rightarrow$ en una muestra de tamaño n $(x_1, \dots, x_n)$ podemos contar el número de caras (1's) y el número de cruces (0's)

$\Rightarrow O_1 \sim Bin(n,p)$, $O_2 \sim Bin(n, 1-p)$ $\qquad O_2 = n - O_1$

$E_1 = np$, $E_2 = n(1-p)$

$$T = \frac{(O_1 - np)^2}{np} + \frac{(n - O_1 - n(1-p))^2}{n(1-p)} = \frac{(O_1 - np)^2}{np} + \frac{(O_1 - np)^2}{n(1-p)} =$$

$$= \frac{(O_1 - np)^2 (1 - p + p)}{np(1-p)} = \frac{(O_1 - np)^2}{np(1-p)} = \Bigg(\underbrace{\frac{O_1 - np}{\sqrt{np(1-p)}}}_{\overset{\shortparallel}{Z} \sim N(0,1)}\Bigg)^2$$

porque:

$$\mathbb{E}(Bin(n,p)) = np$$

$$\mathbb{V}(Bin(n,p)) = np(1-p)$$

Por el TCL:

$$T \xrightarrow{d} Z^2 \sim \chi^2_1$$

con $Z \sim N(0,1)$

$\boxed{2.}$ $(x_1, \dots, x_n)$ ceros y unos $\begin{cases} n^{\text{o}} \text{ unos} = \sum_i x_i = n\bar{x} \\ n^{\text{o}} \text{ ceros} = n - \sum_i x_i = n(1-\bar{x}) \end{cases}$

$X \sim Ber(p)$

	Obs	Esp
#1's	$n\bar{x}$	np
#0's	$n(1-\bar{x})$	$n(1-p)$

$$b = \frac{(n\bar{x} - np)^2}{np} + \frac{(n(1-\bar{x}) - n(1-p))^2}{n(1-p)} =$$

$$= \frac{n^2(\bar{x}-p)^2}{np} + \frac{n^2(1-\bar{x}-(1-p))^2}{n(1-p)} =$$

$$= \frac{n(\bar{x}-p)^2}{p} + \frac{n(\bar{x}-p)^2}{(1-p)} = \frac{n(\bar{x}-p)^2(1-p+p)}{p(1-p)} = \boxed{\frac{n(\bar{x}-p)^2}{p(1-p)}}$$

Entonces, el test sería comparar $\dfrac{n(\bar{x}-p)^2}{p(1-p)}$ con el percentil de la χ^2_1 con nivel de signif. α.

Rechazamos H_0 si $b = \dfrac{n(\bar{x}-p)^2}{p(1-p)} > \chi^2_{\{1;\alpha\}}$

A partir de ahora, hasta los ejercicios de K-S, se realizan en una hoja de Excel. Aquí mostraremos los razonamientos para el grado de libertad.

$\boxed{3.}$ Sería bueno si la distribución fuese uniforme:

$$\text{valor esperado} = \frac{n}{\#\text{valores\_posibles}}$$

En nuestro caso, $\#\text{valores\_posibles} = 10$, $n = 200$

Grados de libertad: $10 - 1 = 9$

$\boxed{4.}$

Rango	Esperados
<-2	$n \cdot \Phi(-2)$
$[-2,-1]$	$n \cdot [\Phi(-1) - \Phi(-2)]$
$(-1, 0]$	$n \cdot [\Phi(0) - \Phi(-1)]$
$(0, 1]$	$n \cdot [\Phi(1) - \Phi(0)]$
$(1, 2]$	$n \cdot [\Phi(2) - \Phi(1)]$
> 2	$n \cdot (1 - \Phi(2))$

Grados de libertad:
$6 - 1 = 5$

$\boxed{5.}$ Esperados en cada celda: $\dfrac{n}{\#\text{celdas}} = \dfrac{60}{9}$

Grados de libertad: $9 - 1 = 8$

$\boxed{6.}$ Estimación por máx. verosimilitud: $\hat{p} = 0'9129$ (solver Excel)

grados de libertad $= 4 - 1 \underline{- 1} = 2$

porque necesitamos estimar p

A mano:

$$\text{VERO}(p) = \left(\frac{p}{2}\right)^{442} \cdot \left(\frac{p^2}{2} + p(1-p)\right)^{514} \cdot \left(\frac{1-p}{2}\right)^{38} \cdot \left(\frac{(1-p)^2}{2}\right)^{6}$$

$$\log \text{VERO}(p) = 442\log\left(\frac{p}{2}\right) + 514\log\Bigg(\underbrace{p - \frac{p^2}{2}}_{=\frac{2p-p^2}{2}}\Bigg) + 38\log\left(\frac{1-p}{2}\right) + 6\log\left(\frac{(1-p)^2}{2}\right)$$

$$\frac{d\log \text{VERO}}{dp}(p) = \frac{442}{p} + 514\,\frac{2(1-p)}{p(2-p)} + \frac{-38}{1-p} + \frac{-12}{1-p} = 0$$

$$\Longleftrightarrow 442(2-p)(1-p) + 1024(1-p)^2 - 50\,p(2-p) = 0 \Longleftrightarrow$$

$$\Longleftrightarrow 884 - 884p - 442p + 442p^2 + 1024 - 2048p + 1024p^2 - 100p + 50p^2 = 0$$

$$\Longleftrightarrow 1516\,p^2 - 3474\,p + 1908 = 0 \Longleftrightarrow$$

$$\Longleftrightarrow p = \frac{3474 \pm \sqrt{(3474)^2 - 4.1516.1908}}{2.1516} = \begin{cases} p = 0'9129 \\ p = 1'3787 \text{ (no válida)} \end{cases}$$

$$\Longrightarrow \boxed{\hat{p} = 91'29\,\%}$$

7.

a) Consideramos H_0: "la prob. de asesinato es igual todos los días"

Grados de libertad: $7 - 1 = 6$

b) Hacemos dos test chi-cuadrado:

1. De lunes a viernes (grados de libertad: $5-1 = 4$)
2. Sábado y domingo, e.d., fin de semana (g.l. $= 2-1 = 1$)

→ Buscamos aceptar ambos para aceptar H_0.

$\boxed{8.}$ Grados de libertad: $4-1=3$

$\boxed{9.}$ Grados de libertad: $2.3-1-(2-1)-(3-1)=2$

$\boxed{10.}$ No se estiman las probabilidades de pertenecer a PU, PT ó CD:

Grados de libertad: $3.3-1-(3-1)=6$

$\boxed{11.}$ No se estiman el nº de zonas por monte.

Grados de libertad: $2.3-1-(3-1)=3$

$\boxed{12.}$ El tamaño de la muestra no se estima.

Grados de libertad: $2.3-1-(2-1)=4$

$\boxed{13.}$ a) Grados libertad: $4-1=3$

b) Grados libertad: $4.2-1-(4-1)-(3-1)=3$

$\boxed{14.}$ Grados de libertad: $3.3-1-(3-1)-(3-1)=4$

$\boxed{15.}$ Grados de libertad: $2.3-1-(2-1)-(3-1)=2$

$\boxed{16.}$ a) Grados de libertad: $2.2-1-(2-1)=2$

b) Calculamos p-valor = 1 − DISTR.CHICUAD(b; 2; TRUE)

donde b se ha calculado en a).

$\Rightarrow$ α para los que se rechazaría: $\alpha \in [\text{p-valor}, 1]$

$\boxed{17.}$ $n=1$, X, F_X continua

Recordemos que, por un tma. visto en clase, la distr. Δ_n es la misma para toda v.a. X continua. En particular, $\Delta_n^X = \Delta_n^U$, con $U \sim \text{UNIF}[0,1]$.

$$\Delta_1 = \max\{|F(X_1)|, |1-F(X_1)|\}$$

Tomamos $F^X = F^{\text{UNIF}[0,1]}$:

$$Z := \max\{F(X_1), 1-F(X_1)\}$$

Hay que estudiar la distr. de $Z = \max\{U, 1-U\}$

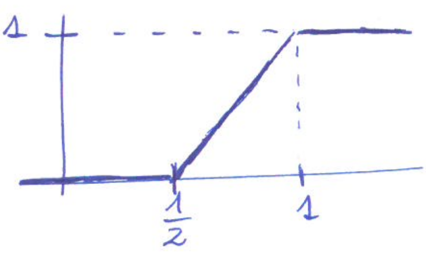
$U \sim \text{UNIF}[0,1]$

Z toma valores en $\left[\frac{1}{2}, 1\right]$:

$$z_0 \in (0,1) \Rightarrow \mathbb{P}(Z \leq z_0) = \mathbb{P}(\max\{U, 1-U\} \leq z_0) =$$

$$= \mathbb{P}(U \leq z_0, 1-U \leq z_0) = \mathbb{P}(1-z_0 \leq U \leq z_0) =$$

$$= \begin{cases} 0 & \text{si } z < \frac{1}{2} \\ 2z-1 & \text{si } z \geqslant \frac{1}{2} \end{cases}$$

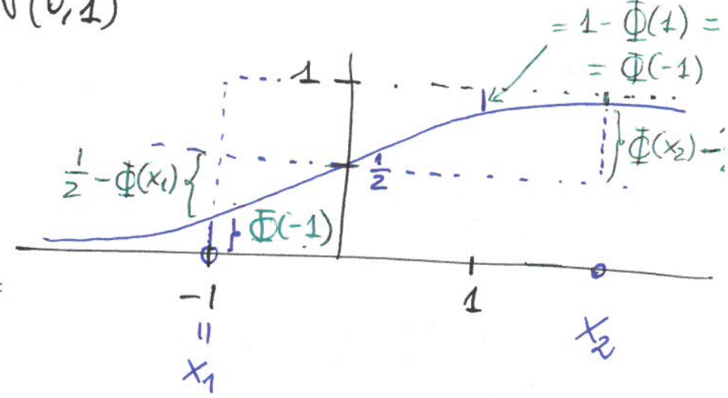


$$\Rightarrow Z \sim \text{UNIF}\left[\frac{1}{2}, 1\right]$$

$\boxed{18.}$ Muestra (x_1, x_2) con $x_1 = -1$ y $x_2 > 1$ desconocido

H_0: muestra tamaño 2 de una $\mathcal{N}(0,1)$

$\delta_n = 0'4192$, ¿x_2?



$$\Delta_2 = \max \left\{ \begin{array}{l} \max\left\{\Phi(-1),\ \frac{1}{2} - \Phi(-1)\right\}, \\ \max\left\{\Phi(x_2) - \frac{1}{2},\ 1 - \Phi(x_2)\right\} \end{array} \right\} =$$

$$= \max \left\{ \begin{array}{l} \max\left\{1 - \Phi(1),\ \frac{1}{2} - 1 + \Phi(1)\right\} \\ \max\left\{\Phi(x_2) - \frac{1}{2},\ 1 - \Phi(x_2)\right\} \end{array} \right\} =$$

$$= \max \left\{ \begin{array}{l} \max\left\{0'1586,\ 0'3413\right\} \\ \max\left\{\Phi(x_2) - \frac{1}{2},\ 1 - \Phi(x_2)\right\} \end{array} \right\}$$

no coinciden con $0'4192$ $\Rightarrow$ no son el máximo

por simetría y sabiendo que Φ es creciente $\Rightarrow \Phi(x_2) - \frac{1}{2} > 1 - \Phi(x_2)$

$$\Rightarrow \Phi(x_2) - \frac{1}{2} = 0'4192 \Rightarrow x_2 = \Phi^{-1}(0'9192) \boxed{= 1'399}$$

$\boxed{19.}$ Sabemos que: $\sqrt{n}\,\Delta_n^+ \xrightarrow[n\to\infty]{d} Z$ con $F_Z(z) = 1 - e^{-2z^2}$ $(z > 0)$

¿convergencia en distribución de $4n(\Delta_n^+)^2$? Sea $W_n = 4n(\Delta_n^+)^2$

todo positivo

$$\omega > 0,\quad \mathbb{P}(W_n \le \omega) = \mathbb{P}\left(4n(\Delta_n^+)^2 \le \omega\right) = \mathbb{P}\left(2\sqrt{n}\,\Delta_n^+ \le \sqrt{\omega}\right) =$$

$$= \mathbb{P}\left(\sqrt{n}\,\Delta_n^+ \le \frac{\sqrt{\omega}}{2}\right) \xrightarrow[n\to\infty]{d} 1 - e^{-2\cdot\frac{\omega}{4}} = \underbrace{1 - e^{-\frac{\omega}{2}}}_{\text{FUNCIÓN DE DISTRIBUCIÓN EXP}(1/2)}$$

$\boxed{20.}$ $\Delta_n = \sup_{t\in\mathbb{R}} |F_n(t) - F(t)| = \max_{1\le i\le n} \left\{ \max\left\{ \left|\frac{i}{n} - F(x_i)\right|, \left|\frac{i-1}{n} - F(x_i)\right| \right\} \right\}$

$\sqrt{n}\,\Delta_n \xrightarrow[n\to\infty]{d}$ cierta variable de K-S.

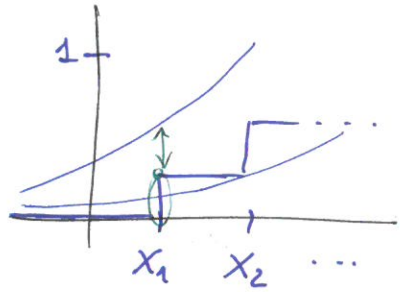
$\Delta_n^+ = \sup_{t\in\mathbb{R}} \left(F_n(t) - F(t)\right)^+$; $\sqrt{n}\,\Delta_n^+ \xrightarrow[n\to\infty]{d} 1 - e^{-2x^2}$, $x>0$

$\Delta_n^- = \sup_{t\in\mathbb{R}} \left(F_n(t) - F(t)\right)^-$

¿Cómo calcular Δ_n^+ para una muestra específica $(x_1, \dots, x_n)$?

$\Delta_n^+ = \max_{1\le i\le n} \left(\frac{i}{n} - F(x_i)\right)^+$

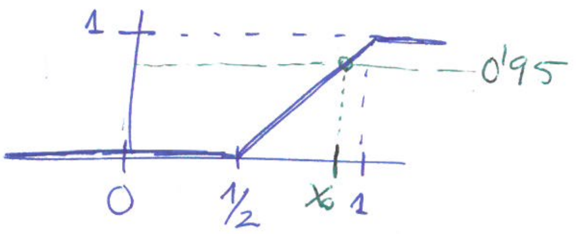
$\sup_{t\le x_1} \left(F_n(t) - F(t)\right)^+ = \left(\frac{1}{n} - F(x_1)\right)^+$

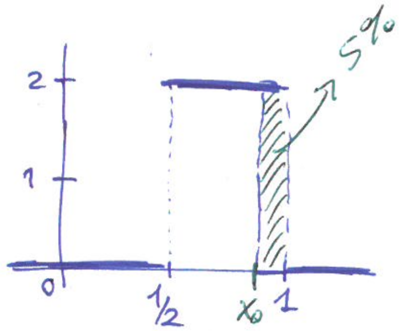


$\boxed{21.}$ Muestra única x H_0: x procede de una $N(0,1)$

Recordemos por el ejercicio 17, K-S de una muestra es

$\Delta_1 \sim \text{UNIF}\left[\frac{1}{2}, 1\right]$





$2x_0 - 1 = 0'95 \longrightarrow x_0 = \frac{1'95}{2} = 0'975$

Rechazaríamos para $\{x\in\mathbb{R} : \underbrace{\Phi(|x|)} \ge 0'975\} \Rightarrow$

$\Rightarrow \{x\in\mathbb{R} : |x| \ge \underbrace{\Phi^{-1}(0'975)}_{=1'96}\}$ valor absoluto debido a la simetría de la normal

Región de rechazo: $\{x\in\mathbb{R} : |x| \ge 1'96\}$