1. Una empresa de internet mide el tiempo de respuesta, que es el tiempo que tarda el servidor en procesar y responder una solicitud del usuario. Para clasificar la calidad del servicio discretiza esta variable en 5 niveles, aplicando la técnica de *binning*. La misma consiste en dividir el soporte de la variable en intervalos disjuntos y reemplazar el valor original de la variable por un valor representativo del intervalo en el que cae. Si el tiempo de respuesta en segundos, X, una variable con distribución exponencial de parámetro $\lambda = 1/2$, aplicar la discretización partiendo el soporte en 5 intervalos de igual probabilidad, representando los valores de cada intervalo por el valor que acumula la mitad de la probabilidad. Sea Y aleatoria resultante:

a) hallar la función de probabilidad de Y,

b) a qué distribución corresponde?,

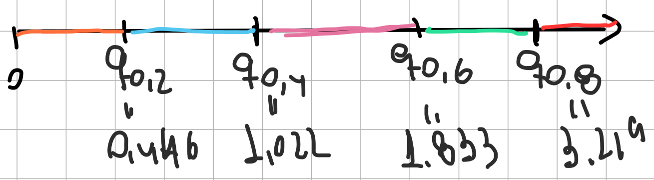
c) hallar la esperanza de Y ,

d) hallar la probabilidad de que Y sea mayor a 1.

$q_{0,2}$ / $P(X < q_{0,2}) = 0,2$

$\underbrace{\quad}_{1 - e^{-q_{0,2}\, 1/2}} = 0,2$

$q_{0,2} = 0,446$



$$Y = \begin{cases} 0.211 & si \ x \in (0, 0.446) \\ 0,713 & si \ x \in (0.446, 1,022) \\ 1,386 & si \ X \in (1,022; 1,833) \\ 2,048 & si \ X \in (1,833; 3,219) \\ 4.605 & si \ X \geq 3,219 \end{cases}$$

a) $p_Y(y) = 0,2 \ \mathbb{1}\{y \in \{0,211; 0,713; 1.386; 2,048; 4,605\}\}$

y	0,21.	0,713	1.386	2,048	4,605
p_Y	1/5	1/5	1/5	1/5	1/5

c) $E[Y] = 0.211 \cdot 1/5 + 0,713 \cdot 1/5 + 0.1386 \cdot 1/5 + 2,048 \cdot 1/5 + 4,605 \cdot 1/5$

$= 1,864$

d $P(Y > 1) = P(Y = 1,386) + P(Y = 2,048) + P(Y = 4,605)$

$= 3/5$

2. Considere un problema de detección, donde se quiere detectar la presencia de una señal de valor constante A desconocido inmersa en ruido gaussiano blanco aditivo $N \sim \mathcal{N}(0, \sigma^2)$ con varianza desconocida. La señal recibida puede representarse como $X = A + N$, de la cual se poseen 15 muestras i.i.d: $X_1, \ldots, X_{15}$.

fijo (no aleatorio)

a) Hallar un intervalo de confianza de nivel 0.95 para la media de la amplitud recibida.

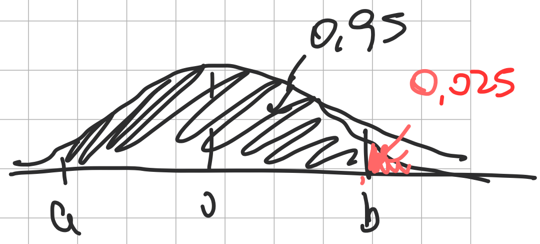
b) Si en las 15 mediciones se observó un valor promedio $\bar{x} = 1.07$ y un desvío estándar muestral $s = 1.87$, determinar si existe evidencia suficiente para asegurar la presencia de señal con 5 % de significación.

$X \sim \mathcal{N}(A, \sigma^2)$ $\quad n = 15$

Busco pivote para media de población normal con var. desconocida

$$U(\underline{X}, A) = \frac{\bar{X} - A}{S/\sqrt{n}} \sim t_{14}$$

$$\mathbb{P}\left(a < \frac{\bar{X} - A}{S/\sqrt{n}} < b\right) = 0.95$$



$$b = t_{14,\, 0.975} = 2{,}145$$

$$a = -b = -2{,}145$$

$$+2{,}145 \frac{S}{\sqrt{n}} + \bar{X} > A > -2{,}145 \frac{S}{\sqrt{n}} + \bar{X}$$

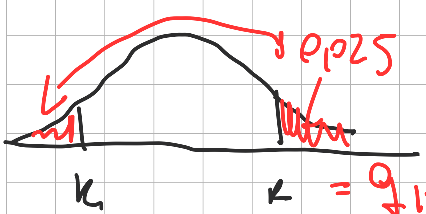
$$\boxed{A \in \bar{X} \pm 2{,}145 \frac{S}{\sqrt{n}}}$$

b) $H_0: A = 0 \qquad H_1: A \neq 0$

$$\delta(\underline{X}) = \mathbb{1}\{\bar{X} > k_1\} + \mathbb{1}\{\bar{X} < k_2\}$$
$$= \mathbb{1}\left\{\left|\frac{\bar{X} - 0}{S/\sqrt{n}}\right| > k\right\}$$

$$\alpha = 0{,}05 = \mathbb{P}_{A=0}(\delta(\underline{X}) = 1) = \mathbb{P}_{A=0}\left(\left|\frac{\bar{X} - 0}{S/\sqrt{n}}\right| > k\right)$$

$$\Rightarrow \delta(\underline{X}) = \mathbb{1}\left\{\left|\frac{\bar{X} - 0}{S/\sqrt{n}}\right| > 2{,}145\right\}$$



$= q_{14,\, 0{,}975} = 2{,}145$

con la muestra $\frac{\bar{x}}{s/\sqrt{15}} = 2{,}216 > 2{,}145$

$\Rightarrow$ Rechazo H_0

(y acepto la presencia de señal)

3. Se desea estimar por regresión lineal el peso de las personas a partir de su altura. Se definen las variables H: *altura de una persona* y W: *peso de una persona*. Se propone el siguiente modelo:

$$W = \beta_0 + \beta_1 H + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0,4).$$

a) Estimar por máxima verosimilitud el valor de los parámetros β_0 y β_1 basandose en una muestra i.i.d. $(H_1, W_1), (H_2, W_2), \cdots, (H_n, W_n)$. *Sugerencia: observar que* $W|H = h \sim \mathcal{N}(\beta_0 + \beta_1 h, 4)$.

b) Comparar con el estimador de mínimos cuadrados, el cual resulta de minimizar $\sum_{i=1}^{n}(w_i - \beta_0 - \beta_1 h_i)^2$.

c) Usar los datos del archivo `Height_Weight.csv` para hallar los valores estimados de β_0 y β_1.

a) $$L(\beta_0,\beta_1) = \prod_{i=1}^{n} f_{W|H=h}(w) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\,2} e^{-\frac{1}{2}\frac{(w_i-(\beta_0+\beta_1 h_i))^2}{4}}$$

$$= \frac{1}{\sqrt{2\pi}\,2} e^{-\frac{\sum(w_i-(\beta_0+\beta_1 h_i))^2}{8}}$$

$$\ell(\beta_0,\beta_1) = \log L(\beta_0,\beta_1) = -\frac{\sum(w_i-(\beta_0+\beta_1 h_i))^2}{8}$$

$$\frac{\partial \ell(\beta_0,\beta_1)}{\partial \beta_0} = \frac{2}{\cancel{8}}\sum(w_i-(\beta_0+\beta_1 h_i)) = 0 \quad (1)$$

$$\frac{\partial \ell(\beta_0,\beta_1)}{\partial \beta_1} = \frac{2}{\cancel{8}}\sum(w_i-(\beta_0+\beta_1 h_i))h_i = 0 \quad (2)$$

(1) $$\sum w_i - \sum \beta_0 - \sum \beta_1 h_i = \frac{\sum w_i - n\beta_0 - \beta_1 \sum h_i}{n} = \frac{0}{n}$$

$$= \bar{w} - \beta_0 - \beta_1 \bar{h} \rightarrow \beta_0 = \bar{w} - \beta_1 \bar{h}$$

(2) $$\frac{\sum w_i h_i - \beta_0 \sum h_i - \beta_1 \sum h_i^2}{n} = \frac{0}{n}$$

$$\frac{\sum w_i h_i}{n} - \beta_0 \bar{h} - \beta_1 \overline{h^2} = 0 \qquad cov(W,H) = E[W \cdot H] - E[W]\,E[H]$$

$$\frac{\sum w_i h_i}{n} - \bar{w}\bar{h} + \underbrace{\beta_1(\bar{h})^2 - \beta_1 \overline{h^2}}_{\beta_1\left((\bar{h})^2 - \overline{h^2}\right)} = 0$$

$$\beta_1 = \frac{-\frac{\sum w_i h_i}{n} + \bar{w}\bar{h}}{(\bar{h})^2 - \overline{h^2}} = \frac{\overbrace{\frac{\sum w_i h_i}{n} - \bar{w}\bar{h}}^{cov(W,H)}}{\underbrace{\overline{h^2} - (\bar{h})^2}_{var(H)}}$$

$$\hat{W} = \frac{cov(W,H)}{var(H)}\left(H - E[H]\right) + E[W]$$