MA34B 14/05/01 **PAUTA CONTROL 2** **N. LACOURLY**

PROBLEMA 1

1.1. $X_i \sim N(1,4) \Rightarrow \bar{x} \sim N(1,\frac{4}{n})$. Si $Z = \frac{\bar{x}-1}{2/\sqrt{n}} \sim N(0,1). \Rightarrow P(|Z| \leq 1.65) = 0.90$

$\Rightarrow P(\bar{x} \in [1-1.65\frac{2}{\sqrt{n}}, 1+1.65\frac{2}{\sqrt{n}}]) = 0.90$

1.2 Tenemos un intervalo a *90%* de confianza y *n=100,*

$$P(\bar{x} \in [1-1.65\frac{2}{\sqrt{n}}, 1+1.65\frac{2}{\sqrt{n}}]) = P(\bar{x} \in [1-1.65\frac{2}{10}, 1+1.65\frac{2}{10}]) = [0.67, 1.33].$$

Para el tamaño de muestra *n*: $P(\bar{x} \in [0.67, 1.33]) = 0.95$

Ahora bien $\sqrt{n}(\bar{x}-1)/2 \sim N(0,1)$, luego $P(\sqrt{n}(\bar{x}-1)/2 \in [-1.96, 1.96]) = 0.95$. o sea

$$P(\bar{x} \in [1-1.96\frac{2}{\sqrt{n}}, 1+1.96\frac{2}{\sqrt{n}}]) = 0.95 \Rightarrow [1-1.96\frac{2}{\sqrt{n}}, 1+1.96\frac{2}{\sqrt{n}}] = [0.67, 1.33]$$

Finalmente: $1+1.96\frac{2}{\sqrt{n}} = 1.33$ y *n>141.*

1.3 Dado que $X \sim N(\mu_1, \sigma^2)$ y $Y \sim N(\mu_2, 2\sigma^2)$ la función de Verosimilitud de las dos muestras es:

$$L = (\frac{1}{\sqrt{2\pi\sigma^2}})^n \exp(-\frac{1}{2\sigma^2}\sum(x_i-\mu_1)^2)(\frac{1}{\sqrt{4\pi\sigma^2}})^m \exp(-\frac{1}{4\sigma^2}\sum(y_i-\mu_2)^2).$$

$$\log(L) = -\frac{1}{2}(n\log(2\pi\sigma^2) + m\log(4\pi\sigma^2)) - \frac{1}{2\sigma^2}\sum(x_i-\mu_1)^2 - \frac{1}{4\sigma^2}\sum(y_i-\mu_2)^2$$

$$\frac{\partial \log(L)}{\partial \mu_1} = \frac{1}{2\sigma^2}\sum(x_i-\mu_1) = 0 \Rightarrow \hat{\mu}_1 = \frac{1}{n}\sum x_i = \bar{x}$$

$$\frac{\partial \log(L)}{\partial \mu_2} = \frac{1}{4\sigma^2}\sum(y_i-\mu_2) = 0 \Rightarrow \hat{\mu}_2 = \frac{1}{m}\sum y_i = \bar{y}$$

$$\frac{\partial \log(L)}{\partial \sigma^2} = -\frac{n+m}{2\sigma^2} + \frac{1}{2\sigma^4}\sum(x_i-\mu_1)^2 + \frac{1}{4\sigma^4}\sum(y_i-\mu_2)^2 = 0 \Rightarrow$$

$\tilde{\sigma}^2 = \frac{1}{n+m}(\sum(x_i-\bar{x})^2 + \frac{1}{2}\sum(y_i-\bar{y})^2)$. Dadas las varianzas muestrales $S_X^2 = \frac{1}{n}\sum_i(x_i-\bar{x})^2$ y $S_Y^2 = \frac{1}{m}\sum_i(y_i-\bar{y})^2$, se tiene: $\tilde{\sigma}^2 = \frac{nS_X^2 + (m/2)S_Y^2}{n+m}$.

Como $E(\frac{n\,S_X^2}{n-1}) = \sigma^2$ y $E(\frac{m\,S_Y^2}{m-1}) = 2\sigma^2$, $E(\tilde{\sigma}^2) = \frac{n+m-2}{n+m}\sigma^2$. es decir que $\tilde{\sigma}^2$ sesgado. Luego se deduce un estimador insesgado para σ^2: $\hat{\sigma}^2 = \frac{nS_X^2 + (m/2)S_Y^2}{n+m-2}$.

1.4 Como $\bar{x} \sim N(\mu_1, \frac{\sigma^2}{n})$ e $\bar{y} \sim N(\mu_2, \frac{2\sigma^2}{m})$, $\bar{x}$ e $\bar{y}$ son independientes,

$$4\bar{x}+\bar{y} \sim N(4\mu_1+\mu_2, \frac{16\sigma^2}{n}+\frac{2\sigma^2}{m})\text{, es decir } \frac{4\bar{x}+\bar{y}-(4\mu_1+\mu_2)}{\sqrt{\frac{16\sigma^2}{n}+\frac{2\sigma^2}{m}}} \sim N(0,1).$$

Por otra parte: $\frac{n\,S_X^2}{\sigma^2} \sim \chi^2_{n-1}$ y $\frac{m\,S_Y^2}{2\sigma^2} \sim \chi^2_{m-1}$.

Como las muestras son independientes, $\frac{n\,S_X^2}{\sigma^2}+\frac{m\,S_Y^2}{2\sigma^2} = (n+m-2)\frac{\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n+m-2}$.

Luego $\frac{4\bar{x}+\bar{y}-(4\mu_1+\mu_2)}{\sqrt{(\frac{16}{n}+\frac{2}{m})\hat{\sigma}^2}} \sim t_{n+m-2}$

Sea $A=\hat{\sigma}\sqrt{\frac{16}{n}+\frac{2}{m}}$**,** entonces el intervalo buscado es de la forma

$[\,[4\bar{x}+\bar{y}-t^{\alpha}_{n+m-2}\text{ A}, 4\bar{x}+\bar{y}+t^{\alpha}_{n+m-2}\text{ A}]$

PROBLEMA 2

2.1 La región crítica más potente del test esta dada por el lema de N.P.: $P(\bar{x} \le c \mid \mu = 20) = 0.01$.

Como $\bar{x} \sim N(\mu, \frac{1}{4})$ => $P(N(0,1) \le 2(c-20)) = 0.01$ => $2(c-20) = -3.09$ => $c = 18.455$.

La región crítica es entonces: $\boxed{\bar{x} \le 18.455}$.

2.2 El error de tipo II es $\beta = P(\bar{x} > 1.8.455 \mid \mu = 18) = P(N(0,1) > 2(18.455-18)) = 0.4637$ =>
$\boxed{\beta = 0.4637}$

2.3
$\alpha = P(\bar{x} \le 18.455 \mid \mu = 20) = P(\sqrt{n}(\bar{x}-20)/3 \le \sqrt{n}(18.455-20)/3) = P(N(0,1) \le -3.60) = 0.0$
$\beta = P(\bar{x} > 18.455 \mid \mu = 18) = P(N(0,1) > 7(18.455-18)/3) = P(N(0,1) > 1.0617 = 0.1442$
$\boxed{\alpha = 0.0 \text{ y } \beta = 0.1442}$

2.4 Para $n=49$ y $\alpha = 0.01$ se obtiene una región crítica:
$\alpha = P(\bar{x} \le c \mid \mu = 20) = P(7(\bar{x}-20)/3 \le 7(c-20)/3) = 0.01$. Es decir $7(c-20)/3 = -2.326$.
Finalmente $c = 19.003$. En el punto 2.2 $\beta = 0.4637$. Si $= 49$ y $\alpha = 0.01$, el error de tipo II se escribe:
$\beta = P(\bar{x} > 19.00 \mid \mu) = P(N(0,1) > 7(19.00-\mu)/3) = 0.4637$. Es decir que hay que cambiar la hipótesis alternativa por: $\boxed{\mu = 18.80}$.

PROBLEMA 3

3.1 $X \sim N(\theta,1)$ => $\bar{x} \sim N(\theta,\frac{1}{n})$. En el caso de estas hipótesis simples, el lema de Neyman Pearson nos lleva a una región crítica de la forma $\bar{x} \geq c$ en donde c esta determinado por el error de tipo I:
$P(\bar{x} \geq c \mid \theta = 1) = 0.05$ => $P(\sqrt{n}(\bar{x}-1) \geq \sqrt{n}(c-1)) = 0.05$ con $\sqrt{n}(\bar{x}-1) \sim N(0,1)$. Las tablas nos dan que $\sqrt{n}(c-1) = 1.65$ => $c = 1+\frac{1.65}{\sqrt{n}} = 1.165$ => $\boxed{\text{se rechaza } H_o : \theta = 1}$ dado que $\bar{x} = 1.2$.

3.2 En el caso de estas nuevas hipótesis simples, el lema de Neyman Pearson nos lleva a una región crítica de la forma $\bar{x} \geq c'$ en donde c' esta determinado por el error de tipo I:
$P(\bar{x} \leq c' \mid \theta = 1.4) = 0.05$ => $P(\sqrt{n}(\bar{x}-1.4) \leq \sqrt{n}(c'-1.4)) = 0.05$ con $\sqrt{n}(\bar{x}-1.4) \sim N(0,1)$.
Las tablas nos dan que $\sqrt{n}(c'-1.4) = -1.65$ => $c' = 1.4-\frac{1.65}{\sqrt{n}} = 1.235$ => $\boxed{\text{se rechaza } H_o' : \theta = 1.4}$
Observamos que en 3.1 y en 3.2 los resultados son opuestos: se rechaza $\theta = 1$ en 3.1 y se rechaza $\theta = 1.4$ en 3.2.
El error de tipo II en 3.1 es igual a:
$P(\bar{x} < 1.165 \mid \theta = 1.4) = P(\sqrt{n}(\bar{x}-1.4) < \sqrt{n}(1.165-1.4)) = P(N(0,1) < -2.35) = 0.0094$.
El error de tipo II en 3.2 es igual a:
$P(\bar{x} > 1.235 \mid \theta = 1) = P(\sqrt{n}(\bar{x}-1) > \sqrt{n}(1.235-1)) = P(N(0,1) > 2.35) = 0.0094$

3.3 Las conclusiones encontradas en 3.1 y 3.2 son opuestas. Para eliminar la contradicción basta tomar n de manera que la región crítica del test de 3.1 sea $\bar{x} \geq 1.2$ y la región crítica del test de 3.2 sea $\bar{x} \leq 1.2$.
Entonces se tiene: $P(\bar{x} \geq 1.2 \mid \theta = 1) = 0.05$ =>
$P(\sqrt{n}(\bar{x}-1) \geq \sqrt{n}(1.2-1)) = P(N(0,1) \geq 0.2\sqrt{n}) = 0.05$ => $0.2\sqrt{n} = 1.65$ => $n=68$.

3.4 Para el test de 3.1:
$P(\bar{x} \geq c \mid \theta = 1) = P(N(0,1) \geq 10(c-1)) = 0.023$ => $10(c-1) = 2$ => $c = 1.2$
Para el test de 3.2
$P(\bar{x} \leq c' \mid \theta = 1.4) = P(N(0,1) \leq 10(c'-1.4)) = 0.023$ => $10(c'-1.4) = -2$ => $c = 1.2$
=> $\boxed{\text{No hay contradicción.}}$