EYP1113 – Pauta Examen Parte II
2020 - 01

Pregunta 1.1

Considere una muestra aleatoria de tamaño n proveniente de una población X cuyo comportamiento probabilístico esta dada por la siguiente función de densidad

$$f(x) = \frac{x^{-(\alpha+1)} e^{-1/(\beta x)}}{\Gamma(\alpha)\beta^{\alpha}}$$

con $x > 0$, $\beta > 0$ y $\alpha > 2$.

Si el valor esperado y varianza de este modelo son:

$$\mu_X = \frac{1}{\beta(\alpha-1)} \qquad \sigma_X^2 = \frac{1}{\beta^2(\alpha-1)^2(\alpha-2)}$$

Entonces el estimador de momento de α es:

$$\frac{1}{\beta(\alpha-1)} = \overline{X} \quad ; \quad \frac{1}{\beta^2(1-\alpha)(2-\alpha)} = \overline{X^2}$$

$$\longrightarrow \quad \hat{\alpha} = \frac{(\overline{X})^2 - 2\,\overline{X^2}}{(\overline{X})^2 - \overline{X^2}}$$

Pregunta 1.2

Considere una muestra aleatoria de tamaño n proveniente de una población X cuyo comportamiento probabilístico esta dada por la siguiente función de densidad

$$f(x) = \frac{x^{-(\alpha+1)}\ \beta^{\alpha}\ e^{-\beta/(x)}}{\Gamma(\alpha)}$$

con $x > 0$, $\beta > 0$ y $\alpha > 2$.

Si el valor esperado y varianza de este modelo son:

$$\mu_X = \frac{\beta}{(\alpha-1)} \qquad \sigma_X^2 = \frac{\beta^2}{(\alpha-1)^2(\alpha-2)}$$

Entonces el estimador de momento de α es:

$$\frac{\beta}{(\alpha-1)} = \overline{X} \quad ; \quad \frac{\beta^2}{(1-\alpha)(2-\alpha)} = \overline{X^2}$$

$$\longrightarrow \quad \hat{\alpha} = \frac{(\overline{X})^2 - 2\,\overline{X^2}}{(\overline{X})^2 - \overline{X^2}}$$

Pregunta 2.1

Los tiempos de permanencia en las UCI ha sido un tema importante en las últimas semanas, ya que podrían implicar un colapso en la atención de pacientes críticos. Al modelar estos tiempos por una Exponencial(ν), tenemos que la distribución asintótica del EMV es

$$\hat{\nu} \overset{\text{aprox}}{\sim} \text{Normal}\left(\nu,\ \frac{\nu}{\sqrt{n}}\right)$$

Por lo tanto la varianza asintótica del EMV de la $P(T > t)$ es:

$$g(v) = e^{-vt} \longrightarrow g'(v) = e^{-vt} \cdot (-t)$$

$$I(v) = \frac{n}{v^2}$$

$$CCR(g(v)) = \frac{(g'(v))^2}{I(v)} = \frac{(vt)^2 e^{-2vt}}{n}$$

Pregunta 2.2

Los tiempos de permanencia en las UCI ha sido un tema importante en las últimas semanas, ya que podrían implicar un colapso en la atención de pacientes críticos. Al modelar estos tiempos por una Exponencial(ν), tenemos que la distribución asintótica del EMV es

$$\hat{\nu} \overset{\text{aprox}}{\sim} \text{Normal}\left(\nu,\ \frac{\nu}{\sqrt{n}}\right)$$

Por lo tanto la varianza asintótica del EMV de la mediana es:

$$g(\nu) = \frac{\ln(2)}{\nu} \longrightarrow g'(\nu) = -\frac{\ln(2)}{\nu^2}$$

$$I(\nu) = \frac{n}{\nu^2}$$

$$CCR(g(\nu)) = \frac{(g'(\nu))^2}{I(\nu)} = \frac{(\ln(2))^2}{n\,\nu^2}$$

La comuna de Santiago enfrenta la pandemia día a día, incrementando la toma de PCR.

Usted logra acceder a una base de pruebas PCR que fueron tomados durante un par de días [Pinche para Descargar]. En ella se dispone de información respecto a los pacientes que concurren a un CESFAM a tomarse el test PCR:

- CASO: (ID)
- EXTRANJERO (0 = No, 1 = Si)
- PCR: (0 = Negativo, 1 = Positivo)
- EDAD (En años)

A continuación en base a esta información responda las siguientes dos preguntas.

Recuerde DESCARGAR los datos y luego presione el botón SIGUIENTE para poder ver las preguntas.

```
Data = read.table("PCR.txt", header = T)
```

Pregunta 3.1

Para la siguiente afirmación lleve a cabo el test de hipótesis correspondiente y entregue el valor-p.

¿Existe evidencia que permita afirmar que más de 1/3 de los pacientes que concurren son extranjeros residentes?

```
n = 112
p.hat = 49/n
p0 = 1/3

## A mano
Zo = (p.hat-p0)/sqrt(p0*(1-p0)/n)
1-pnorm(Zo)
[1] 0.009679734

## Utilizando z.test()
z.test(x = p.hat, mu = p0, sd = sqrt(p0*(1-p0)/n), alternative =
 "greater")$p.value
[1] 0.009679734

## Utilizando prop.test()
prop.test(x = 49, n = n, p = p0, correct = F, alternative = "great")$p.value
[1] 0.009679734
```

Pregunta 3.2

Para la siguiente afirmación lleve a cabo el test de hipótesis correspondiente y entregue el valor-p. ¿Existe evidencia que permita afirmar que la mayoría de los pacientes que concurren no son extranjeros?

```
n = 112
p.hat = 63/n
p0 = 1/2

## A mano
Zo = (p.hat-p0)/sqrt(p0*(1-p0)/n)
1-pnorm(Zo)
[1] 0.09293837

## Utilizando z.test()
z.test(x = p.hat, mu = p0, sd = sqrt(p0*(1-p0)/n), alternative =
 "greater")$p.value
[1] 0.09293837

## Utilizando prop.test()
prop.test(x = 63, n = n, p = p0, correct = F, alternative = "great")$p.value
[1] 0.09293837
```

Pregunta 3.3

Para la siguiente afirmación lleve a cabo el test de hipótesis correspondiente y entregue el valor-p. ¿Es posible afirmar que la positividad es inferior a un 30%?

```
n = 112
p.hat = 25/n
p0 = 0.3

## A mano
Zo = (p.hat-p0)/sqrt(p0*(1-p0)/n)
pnorm(Zo)
[1] 0.03809032

## Utilizando z.test()
z.test(x = p.hat, mu = p0, sd = sqrt(p0*(1-p0)/n), alternative =
 "less")$p.value
[1] 0.03809032

## Utilizando prop.test()
prop.test(x = 25, n = n, p = p0, correct = F, alternative = "less")$p.value
[1] 0.03809032
```

Pregunta 4.1

Para la siguiente afirmación lleve a cabo el test de hipótesis correspondiente y entregue el valor-p.

Suponiendo que no es posible asumir varianzas iguales, ¿existe evidencia que permita afirmar que los extranjeros que concurren al CESFAM son de mayor edad en comparación con los no extranjeros?

```
X = Data$EDAD[Data$EXTRANJERO == 1]
Y = Data$EDAD[Data$EXTRANJERO == 0]

## Utilizando z.test()
t.test(x = X, y = Y, var.equal = F, alternative = "great")$p.value
[1] 0.001155711
```

Pregunta 4.2

Para la siguiente afirmación lleve a cabo el test de hipótesis correspondiente y entregue el valor-p.

Suponiendo que es posible asumir varianzas iguales, ¿existe evidencia que permita afirmar que los pacientes que presentan un PCR positivo son de mayor edad en comparación con los pacientes que presentan un PCR negativo?

```
X = Data$EDAD[Data$PCR == 1]
Y = Data$EDAD[Data$PCR == 0]

## Utilizando z.test()
t.test(x = X, y = Y, var.equal = T, alternative = "great")$p.value
[1] 0.002619111
```

Pregunta 5.1

En políticas públicas de salud, el estado nutricional de las personas es muy importante. Le encargan realizar un análisis y antes de comenzar desea testear si el IMC de su muestra distribuye Log-Normal.

A partir de los EMV se estimaron las probabilidades para los siguientes intervalos:

```
------------------------------
        Observado Probabilidad
------------------------------
  <=20          7       0.0577
(20-25]        36       0.2875
(25-30]        45       0.3680
(30-35]        23       0.2023
  >35          11       0.0845
------------------------------
```

Entregue el valor-p de la prueba estadística correspondiente.

```
O = c(7,36,45,23,11)
p = c(0.0577, 0.2875, 0.3680, 0.2023, 0.0845)
X2 = chisq.test(x = O, p = p)$statistic
1-pchisq(X2, df = 5 - 1 - 2)
[1] 0.9113706
```

Pregunta 5.2

En políticas públicas de salud, el estado nutricional de las personas es muy importante. Le encargan realizar un análisis y antes de comenzar desea testear si el IMC de su muestra distribuye Log-Normal.

A partir de los EMV se estimaron las probabilidades para los siguientes intervalos:

```
------------------------------
        Observado Probabilidad
------------------------------
  <=20          7       0.0827
(20-25]        36       0.2410
(25-30]        45       0.3580
(30-35]        23       0.2379
  >35          11       0.0804
------------------------------
```

Entregue el valor-p de la prueba estadística correspondiente.

```
O = c(7,36,45,23,11)
p = c(0.0827, 0.2410, 0.3580, 0.2379, 0.0804)
X2 = chisq.test(x = O, p = p)$statistic
1-pchisq(X2, df = 5 - 1 - 2)
[1] 0.1450296
```

Pregunta 6

Suponga que la prueba de bondad de ajuste anterior para un modelo Log-Normal y Normal entrego los siguientes resultados:

```
            statistic   p-value
Log-Normal       0.20     0.905
Normal           3.80     0.150
```

Realice la prueba de hipótesis en base a los EMV correspondiente para la siguiente afirmación: "Más de la mitad de la poblacional adulta tiene un indice de masa corporal mayor a 25%"

```
  sum(X) sum(X^2) sum(log(X)) sum(log(X)^2)   N
3350.035 95516.78    401.9215      1328.496 122
```

Apoya la afirmación: [respuesta]

Utilice un nivel de significancia del 1% en su decisión

```
## Test Mediana Log-Normal

## Ho: exp(lambda) = 25 vs Ha: exp(lambda) > 25
mediana0 = 25
mediana.hat = exp(401.9215/122)
zeta.hat = sqrt(1328.496/122 - (401.9215/122)^2)
Zo = (mediana.hat - mediana0)/sqrt(25^2*zeta.hat^2/n)
valor.p = 1-pnorm(Zo)
M = cbind(Zo, valor.p)
colnames(M) = c("z.value", "p.value")
M
      z.value      p.value
[1,] 4.378845 5.965509e-06

## Ho: lambda = log(25) vs Ha:     lambda  > log(25)
lambda0 = log(25)
lambda.hat = 401.9215/122
zeta.hat = sqrt(1328.496/122 - (401.9215/122)^2)
Zo = (lambda.hat - lambda0)/sqrt(zeta.hat^2/n)
valor.p = 1-pnorm(Zo)
M = cbind(Zo, valor.p)
colnames(M) = c("z.value", "p.value")
M
      z.value      p.value
[1,] 4.215489 1.246187e-05
```

Existe suficiente evidencia al 1% para rechazar Ho y por lo tanto podemos apoyar que más de la mitad de la poblacional adulta tiene un indice de masa corporal mayor a 25%.

Pregunta 7

El siguiente modelo de regresión lineal fue creado para predecir la velocidad de un auto en base a su distancia de frenado.

Determine el valor del coeficiente de determinación ajustado de este modelo.

```
lm(formula = Velocidad ~ Distancia, data = data)

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    7.93930    1.00397   7.908 6.31e-09
Distancia      0.14797    0.02889   5.122 1.51e-05

Residual standard error: 2.761 on 31 degrees of freedom
Multiple R-squared: 0.4584, Adjusted R-squared: XXX
F-statistic: XXX on XXX and 31 DF, p-value: 1.513e-05

Analysis of Variance Table

Response: Velocidad
          Df Sum Sq Mean Sq F value    Pr(>F)
Distancia XX 199.94      XX  26.237 1.513e-05
Residuals 31     XX      XX
```

Adjusted R-squared:0.4409

$$R^2 = 1 - \frac{SCE}{SCT} \longrightarrow 1 - R^2 = \frac{SCE}{SCT}$$

$$\longrightarrow \frac{(n-1)}{(n-2)}(1-R^2) = \frac{SCE/(n-2)}{SCT/(n-1)}$$

$$\longrightarrow 1 - \frac{(n-1)}{(n-2)}(1-R^2) = 1 - \frac{SCE/(n-2)}{SCT/(n-1)}$$

$$= r^2$$

REMPLAZANDO $R^2 = 0.4584$ y $n = 33$

$\longrightarrow r^2 = 0.440929$ //