Pontificia Universidad Católica de Chile
Facultad de Matemáticas
Departamento de Estadística

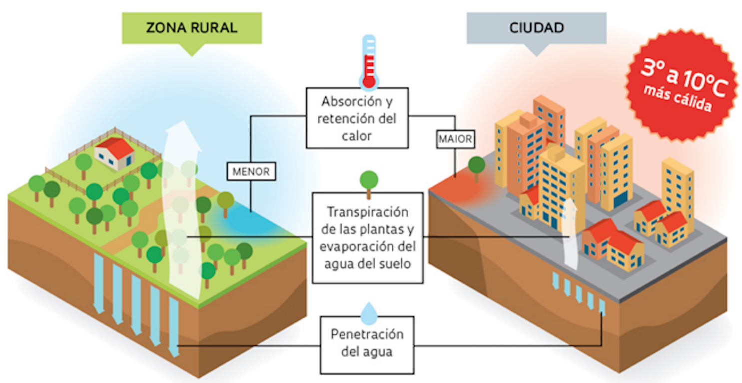
Temporada Académica de Verano 2017

Curso : Probabilidad y Estadística
Sigla : EYP1113
Profesores : Ricardo Aravena C. y Ricardo Olea O.

Pauta Examen

Problema 1

Recientemente se publicó un trabajo referente a las islas de calor que aparecen en Santiago, las cuales son zonas cuyas temperaturas son más alta de la oficial reportada para la ciudad.



Usted, agobiado por las temperaturas nocturnas decide llevar a cabo un experimento que le permita demostrar que las mentadas islas de calor no son ideas afiebradas de investigadores auspiciados por empresas de aire acondicionado, sino que es una realidad con la cual usted debe convivir. Para ello registra durante un mes calendario (30 días) dos zonas específicas de Santiago (su casa que según los expertos se encuentra en una isla de calor, y la universidad, donde abundan las áreas verdes por lo cual la temperatura debiera ser más baja) registrando la temperatura ambiente a las 21.00 hrs. Parte de los resultados se presentan a continuación: En su casa, durante los 15 días en que se superó en más de un grado la temperatura oficial de la ciudad, la temperatura promedio fue 19.5 grados y una desviación estándar igual a 2.4 grados. Por otra parte, en la Universidad solo en nueve ocasiones se superó en más de un grado la temperatura oficial de la ciudad, siendo el promedio en esos días igual a 17.5 grados con una desviación estándar igual a 3.1 grados. ¿Es posible afirmar que su casa se encuentra en una zona de calor? Realice todos los test posibles con nivel de significancia del 5 % y asuma normalidad en las mediciones.

Solución

1) DATOS

		CASA	UNIV
	nº de días	30	30
	+ 1ºC Prom	15	9
Temp	n	15	9
	$\bar{x}$	19,5	17,5
	s	2,4	3,1

a) Test de Proporciones

Sea P_C = proporción de días sobre 1ºC CASA
P_U = ✓ ✓ ✓ ✓ UNIV

0,5 $H_0: P_C = P_U$ vs $H_1: \underbrace{P_C > P_U}_{\text{CASA - ISLA de calor}}$ (o Ha)

Se tiene $\hat{P}_C = 15/30 = 0{,}5$ $\hat{P}_U = 9/30 = 0{,}3$ 0,3

$\bar{P} = \frac{24}{60} = 0{,}4$

$$\text{Test } Z = \frac{0{,}5 - 0{,}3}{\sqrt{0{,}4 \times 0{,}6\left(\frac{1}{30} + \frac{1}{30}\right)}} = 1{,}58 \quad 0{,}5$$

Como $Z_0 = 1{,}58 \not> K_{0,95} = 1{,}645$ 0,2

no hay evidencia contra Ho, es decir la proporción de días con más de 1ºC en mi caso contrastado con la univ NO permite deducir que vivo en una ISLA de CALOR

0,5

b) Veamos si la T° media casa vs univ permite deducir

sea μ_c = Temp media casa - DIAS > 1°C

μ_U = " - univ - " -

0,5 $H_0: \mu_c = \mu_u$ vs $\underbrace{H_1: \mu_c > \mu_u}_{\text{CASA - ISLA de calor}}$ (o H_a)

¿Varianzas?

- Previo $H_0: \sigma_c = \sigma_u$ vs $H_1: \sigma_c \neq \sigma_u$ 0,5

Como $F_c = \underbrace{\dfrac{3{,}1^2}{2{,}4^2} = 1{,}668}_{0{,}5} \not> \underbrace{F_{0{,}975}(8,14) = 3{,}29}_{0{,}5}$

no hay evidencia contra H_0. Puedo Asumir σ's desconocidas pero iguales 0,5

(también podrían concluir $F_c \not< F_{0{,}025}(8,14) \simeq 0, \ldots$

y se hacen $F_c = \left(\dfrac{2{,}4}{3{,}1}\right)^2 = 0{,}599 \not< F_{0{,}025}(14,8) = \dfrac{1}{3{,}29} = 0{,}303$

misma conclusión

Test $t = \dfrac{19{,}5 - 17{,}5}{2{,}676\sqrt{\frac{1}{15} + \frac{1}{9}}} = 1{,}773$ 0,5

Como $t_0 = 1{,}77 > t_{0{,}95}(22) = 1{,}717$ 0,3

se Rechaza H_0. Es decir. ¡VIVO en una ISLA de Calor! 0,5

0,2 $* \; S_c^2 = \dfrac{(15-1) \cdot 2{,}4^2 + (9-1)\, 3{,}1^2}{(15+9-2)} = 7{,}1$

+ 1 Punto Base

Problema 2

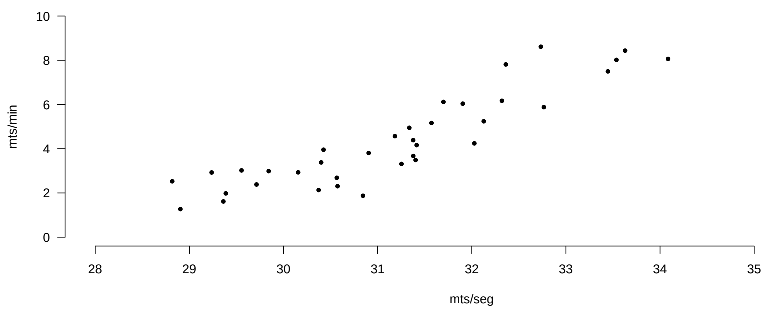
Durante los últimos días se está luchando por apagar los focos de incendio que afectan la zona centro-sur del país. Durante la planificación de la estrategia para extinguir dichos focos, se recibe información del avance del fuego en términos de su velocidad de avance Y (mts/min) en 36 puntos:

Velocidad (mts/min)	≤ 2	$(2-4]$	$(4-6]$	$(6-8]$	> 8
Frecuencia	4	15	8	5	4

```
  mean(Y) mean(Y^2) mean(log(Y)) mean((log(Y))^2)
 4.378002  23.51589     1.358995         2.092202
```

(a) ¿Entre una distribución Log-Normal y Normal, cuál ofrece un mejor ajuste? Use los estimadores de máxima verosimilitud y un $\alpha = 5\,\%$.

Un experto afirma que la velocidad del avance del fuego está relacionada positivamente con la velocidad del viento X en mts/seg. Para demostrar lo anterior obtiene desde el explorador eólico que tiene el (http://walker.dgf.uchile.cl/Explorador/Eolico2/) las velocidades del viento en los 36 puntos y las relaciona con los datos estudiados en (a)



Las correlación lineales obtenidas se presenta a continuación:

```
               Y    log(Y)         X
Y      1.0000000 0.9688312 0.8517558
log(Y) 0.9688312 1.0000000 0.8869140
X      0.8517558 0.8869140 1.0000000

  mean(X) mean(X^2)
 31.18379  974.2888
```

(b) Obtenga la recta de regresión más adecuada, según el resultado obtenido en (a) y determine su coeficiente de determinación.

Solución

(a) Se pide contrastar las siguientes hipótesis

$$\text{H}_0 : Y \sim \text{Normal}(\mu, \sigma) \quad \text{vs} \quad \text{H}_a : Y \not\sim \text{Normal}(\mu, \sigma) \quad \textbf{[0.3 Puntos]}$$

para el cálculo de probabilidades utilizaremos los estimadores máximo verosímiles

$$\textbf{[0.3 Puntos]} \quad \hat{\mu} = \overline{Y} = 4.378002 \quad \text{y} \quad \hat{\sigma} = \sqrt{\overline{Y^2} - \left(\overline{Y}\right)^2} = 2.085423 \quad \textbf{[0.3 Puntos]}$$

Las probabilidades y valores esperados se presentan a continuación

	Observado (O)	Probabilidad	Esperado (E)	(O-E)^2/E
<=2	4	0.1271	4.5756	0.07240916
(2-4]	15	0.3015	10.8540	1.58368491
(4-6]	8	0.3537	12.7332	1.75943064
(6-8]	5	0.1768	6.3648	0.29265319
>8	4	0.0409	1.4724	4.33901233
Total	36	1.0000	36.0000	8.04719024

[0.3 Puntos]

En este caso, se tiene que el estadístico de prueba $X^2 = 8.047$ y al comparar con una $\chi^2(2)$ se tiene que

$$1\,\% < \text{valor-p} < 2.5\,\%$$

Por lo tanto, existe suficiente evidencia para rechazar el ajuste Log-Normal. **[0.3 Puntos]**

Se pide además contrastar las siguientes hipótesis

$$\text{H}_0 : Y \sim \text{Log-Normal}(\lambda,\, \zeta) \quad \text{vs} \quad \text{H}_a : Y \nsim \text{Log-Normal}(\lambda,\, \zeta)$$ **[0.3 Puntos]**

para el cálculo de probabilidades utilizaremos los estimadores máximo verosímiles

[0.3 Puntos] $\hat{\lambda} = \overline{\log(Y)} = 1.358995$ y $\hat{\zeta} = \sqrt{\overline{(\log(Y)^2} - \left(\overline{\log(Y)}\right)^2} = 0.4953126$ **[0.3 Puntos]**

Las probabilidades y valores esperados se presentan a continuación

	Observado (O)	Probabilidad	Esperado (E)	(O-E)^2/E
<=2	4	0.0901	3.2436	0.17639073
(2-4]	15	0.4338	15.6168	0.02436109
(4-6]	8	0.2839	10.2204	0.48238583
(6-8]	5	0.1187	4.2732	0.12361655
>8	4	0.0735	2.6460	0.69286319
Total	36	1.0000	36.0000	1.49961739

[0.3 Puntos]

En este caso, se tiene que el estadístico de prueba $X^2 = 1.4996$ y al comparar con una $\chi^2(2)$ se tiene que

$$10\,\% < \text{valor-p} < 90\,\%$$

Por lo tanto, no existe suficiente evidencia para rechazar el ajuste Log-Normal. **[0.3 Puntos]**

(b) Lo que se pide se obtiene a partir de las siguientes igualdades

$$\hat{\rho} = \hat{\beta}\,\frac{s_X}{s_Y}, \qquad s_X = \sqrt{\frac{n\,\overline{X^2} - n\,\left(\overline{X}\right)^2}{n-1}}, \qquad s_Y = \sqrt{\frac{n\,\overline{Y^2} - n\,\left(\overline{Y}\right)^2}{n-1}}$$

$$\hat{\alpha} = \overline{Y} - \hat{\beta}\,\overline{Y}, \qquad s_{Y|x} = \sqrt{\frac{(n-1)}{(n-2)}\,(s_Y^2 - \hat{\beta}^2\,s_X^2)}, \qquad r^2 = 1 - \frac{s_{Y|x}^2}{s_Y^2}$$

Alternativa 1: $Y \sim$ Log-Normal

```
      Sx        Sy      rho         b         a       Syx        r2
1.383179 0.5023387 0.886914 0.3221066 -8.685509 0.2321424 0.7803405
```

Asignar **[1.0 Puntos]** por b, **[1.0 Puntos]** por a y **[1.0 Puntos]** por r^2

Alternativa 2: $Y \sim$ Normal

```
      Sx       Sy       rho         b         a       Syx        r2
1.383179 2.115005 0.8517558 1.3024110 -36.23610 1.1085860 0.7174141
```

Asignar **[1.0 Puntos]** por b, **[1.0 Puntos]** por a y **[1.0 Puntos]** por r^2

+ 1 Punto Base