UNIVERSIDAD EAFIT

DEPARTAMENTO DE CIENCIAS MATEMÁTICAS
Estadística 2 CM0418
parcial 3 (25%)

Nombre: David Plazas Escudero Código: 201710005101 Nota: 5.0

Profesor: Francisco Zuluaga Grupo: 001 Octubre 25 de 2018

1. En condiciones normales , ¿ El promedio de temperatura corporal es igual para hombres y mujeres?
Investigadores médicos interesados en esta pregunta recolectaron datos de un gran numero de hombres y mujeres, y en la siguiente tabla se presentan muestras aleatorias de los datos sobre temperaturas corporales en grados farenheith

HOMBRES Y_{1i}	MUJERES Y_{2i}
96.9	97.8
97.4	98
97.5	98.2
97.8	98.2
97.8	98.2
97.9	98.6
98.0	98.8
98.6	99.2
98.8	99.4

a) Encuentre un intervalo de confianza del 95% para la diferencia de medias. ¿ Que puede decir con relación a la pregunta planteada incialmente?

$\mu_1 - \mu_2 \xrightarrow{ep} \bar{Y}_1 - \bar{Y}_2 \quad 1-\alpha = 0.95 \quad n_1 = n_2 = 9 \to$ Pequeño

$\bar{Y}_1 = 97.8556 \qquad \bar{Y}_2 = 98.4889$

$S_1 = 0.5833 \qquad S_2 = 0.5487$

$S_1^2 = 0.3402 \qquad S_2^2 = 0.3011$

$$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$$

$$S_p^2 = \frac{(8)(0.5833) + (8)(0.3011)}{8+8-2}$$

$S_p^2 = 0.5053$

$S_p = 0.7109$

$1-\alpha = 0.95$

$\alpha = 0.05$

$\frac{\alpha}{2} = 0.025$

$$[\hat{\theta}_L, \hat{\theta}_u] = \bar{Y}_1 - \bar{Y}_2 \pm t_{\frac{\alpha}{2}, n_1+n_2-2} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$= -0.6333 \pm t_{0.025,14}(0.7109)\sqrt{\frac{2}{9}}$$

$$= -0.6333 \pm 0.7188$$

$$= [-1.3521, 0.0855]$$

1.0

Como el intervalo está tanto positivo como negativo, no se puede afirmar que una de las medias sea mayor.

No se puede afirmar que una de las temperaturas (en promedio), sea mayor.

$$\frac{\sigma_2^2}{\sigma_1^2} \rightarrow \left[\frac{S_2^2}{S_1^2} \frac{1}{\mathbb{F}_{n_2-1, n_1-1, \frac{\alpha}{2}}} , \frac{S_2^2}{S_1^2} \mathbb{F}_{n_1-1, n_2-1, \frac{\alpha}{2}} \right]$$

2. En un estudio sobre mediciones de la presión del compartimiento anterior del musculo(en milímetros de mercurio) se tomaron 10 mediciones para corredores sanos y para ciclistas sanos

1	2
CORREDORES EN ESTADO REPOSO	CICLISTAS EN ESTADO DE REPOSO
MEDIA MUESTRAL = 14.5	MEDIA MUESTRAL = 11.1
VARIANZA MUESTRAL = 15.3664	VARIANZA MUESTRAL =24.5025
MUESTRA = 10	MUESTRA = 10

$$n_1 = n_2 = 10 \rightarrow n_1 - 1 = n_2 - 1 = 9$$

a) Encuentre un intervalo de confianza del 95 % para $\frac{\sigma_2^2}{\sigma_1^2}$

$1 - \alpha = 0.95$

$\alpha = 0.05$

$\frac{\alpha}{2} = 0.025$

$$\frac{S_2^2}{S_1^2} = \frac{15.3664}{24.5025} = 0.6271$$

$$\mathbb{F}_{9,9,0.025} = 4.03$$

$$\left[0.6271 \cdot \frac{1}{4.03} , 0.6271(4.03) \right] = \left[0.1556, 2.5272 \right]$$

UNIVERSIDAD EAFIT
Abierta al mundo

b) ¿ Analizando el intervalo encontrado en el numeral (a) se puede concluir que Los datos dan suficiente evidencia para indicar una variabilidad en presión del compartimento anterior del musculo es mayor para corredeores en estado de reposo que para ciclistas?

Para analizar un intervalo para $\frac{\sigma_2^2}{\sigma_1^2}$, se debe verificar si es completamente menor a 1, o completamente mayor.
En este caso, como el intervalo contiene el 1, no se puede afirmar que la variabilidad en presión del compartimiento anterior del músculo es mayor o menor para corredores o ciclistas; con un nivel de confianza del 95%.

1.0

3. Lord Rayleigh fue uno de los primeros científicos en estudiar la densidad del nitorgeno.En sus estudios observo algo particular.Las densidades del nitrógenos producido a partir de compuestos químicos tendrían a ser menores que las densidades del nitorgeno producido del aire. Las mediciones de Lord Rayleigh se dan en la siguiente tabla.Estas mediciones correponden a la masa de nitrógeno que llena un frasco de volumen especificado a temperatura y presión especificadas.

COMPUESTO QUIMICO (1)	ATMOSFERA (2)
2.30143	2.31017
2.29890	2.30986
2.29816	2.31010
2.30182	2.31001
2.29869	2.31024
2.29940	2.31010
2.29849	2.31028
2.29889	2.30956
2.30057	

$n_1 = 9 \quad n_2 = 8 \qquad 1-\alpha = 0.95 \quad \frac{\alpha}{2} = 0.025$

a) Encuentre un intervalo de confianza del 95% para la diferencia en masa media de nitrógeno por frasco para compuestos químicos y Atmosfera

$\bar{Y}_1 = 2.299594444$
$S_1 = 1.340980156 \times 10^{-3}$
$S_1^2 = 1.798227778 \times 10^{-6}$
$n_1 = 9$
$n_1 - 1 = 8$

$\bar{Y}_2 = 2.31004$
$S_2 = 2.346425853 \times 10^{-4}$
$S_2^2 = 5.505714286 \times 10^{-8}$
$n_2 = 8$
$n_2 - 1 = 7$

$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$
$= 9.84748483 \times 10^{-7}$
$S_p = 9.923447729 \times 10^{-4}$

UNIVERSIDAD EAFIT Abierta al mundo

$$\bar{Y}_1 - \bar{Y}_2 \pm t_{\frac{\alpha}{2}, n_1+n_2-2} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$S_p = 9.923447729 \times 10^{-4}$

$\bar{Y}_1 - \bar{Y}_2 = -0.010445556$ $\quad [\hat{\theta}_L, \hat{\theta}_u] = [-0.011473, -9.41800296 \times 10^{-3}]$

$t_{0.025, 15} = 2.131$

$\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = 0.4859126579$

1.0

b) Se puede concluir que la masa media a partir de compuestos químicos es menor que la obtenida para la atmosfera?

Como el intervalo es completamente negativo, significa que μ_2 es mayor a μ_1, i.e. el promedio es mayor en la atmósfera. Entonces sí, se puede afirmar que la masa media es menor en compuestos químicos. Esto con un nivel de confianza del 95%.

4 Un servicio de cobro de cheques bancarios descubrió que alrededor del 5% de todos los cheques enviados al servicio no tenían fondos.Despues de instituir un sistema de verificación de cheques para reducir sus perdidas , el servicio encontró que de una muestra aleatoria de 1124 que fueron cobrados en efectivo solo 45 cheques carecian de fondos.

$P_0 = 0.05 \quad \hat{P} = \frac{45}{1124} = 0.04 \quad n = 1124$ - Grande.

Se puede concluir que el sistema de verficacion redujo el porcentaje de cheques sin fondos? Utilice un alpha de 0.05

$H_0: P = 0.05$ $\qquad \sigma_{\hat{p}} = \sqrt{\hat{P}(1-\hat{P})/n}$

$H_a: P < 0.05 \rightarrow$ Cola inferior $\longrightarrow RR: \{Z < -Z_\alpha\}$

$Z = \frac{\hat{P} - P_0}{\sigma_{\hat{p}}} \sim N(0,1)$ $\qquad RR: \{Z < -Z_{0.05}\}$.

$RR: \{Z < -1.64\}$.

$$Z = \frac{0.04 - 0.05}{\sqrt{(0.04)(0.96)/1124}} = -1.7108$$

1.0

↓

UNIVERSIDAD EAFIT Abierta al mundo

Como Z está en la región de rechazo (RR), podemos rechazar H_0 ya que la prueba de hipótesis proporcionó evidencia a favor de H_a: el sistema redujo el número de cheques sin fondos.
Esto, con significancia de 0.01; es decir, la probabilidad de encontrar "falsos positivos" es de 0.01.

5. En un estudio relacionado con la polinización controlada de Phlox drummondii,una planta anual que florea en primavera, se concluyo que la falta de agua o nutrientes no influia en los indices de supervivencia de las semillas.El experimento consistía en identificar a las flores que donaban polen con el gnerro masculino y a las flores polinizadas con el genero femenino y clasificarlas en tres grupos: bajo control,con poca agua y con pocos nutrientes. Los datos que se presentan en la siguiente tabla reflejan un aspecto de los hallazgos del experimento:la cantidad de semillas que sobreviven hasta la madurez en cada uno de los tres grupos, que incluyen los progenitores del genero masculino y femenino.

Genero Masculino

tratamiento	n	# de supervivientes
Bajo control	585	543
Poca agua	578	522
Pocos nutrientes	568	510

① — $\hat{P}_1 = \frac{522}{578} = 0.903$

② — $\hat{P}_2 = \frac{510}{568} = 0.898$

Existe evidencia que la proporcion de supervivientes del grupo de semillas con poca agua es mayor grupo de semillas con pocos nutrientes, en el caso de los progenitores del genero masculino? Utilice un nivel de significancia de 0.01 $\alpha = 0.01$

$H_0: P_1 - P_2 = D_0$ $\quad Z = \frac{\hat{P}_1 - \hat{P}_2 - D_0}{\sigma_{\hat{P}_1 - \hat{P}_2}}$ $\quad RR = \{Z > Z_\alpha\} = \{Z > 2.33\}$

$H_a: P_1 - P_2 > D_0$

Cola Superior.

1.0

UNIVERSIDAD EAFIT Abierta al mundo

Para $D_0 = 0$

$$Z = \frac{\hat{P}_1 - \hat{P}_2}{\sigma_{\hat{p}_1 - \hat{p}_2}}$$

$$\sigma_{\hat{p}_1 - \hat{p}_2} = \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

$$= 0.0177$$

$\hat{P}_2 = 0.898$, $\hat{P}_1 = 0.903$

$RR = \{Z > 2.33\}$

$$Z = \frac{0.903 - 0.898}{0.0177} = 0.2824.$$

Como Z no está en la región de rechazo, no se pudo encontrar evidencia para la hipótesis alterna. Luego, no se rechaza H_0. En otras palabras, no se encontró evidencia para afirmar que la proporción de supervivencia de semillas con poca agua es mayor al grupo de semillas con pocos nutrientes.

Nota : Cada punto tiene un valor de 1.0

ESTADISTICO DE PRUEBA MUESTRAS GRANDES:

$$\mathbf{Z} = \frac{\hat{\theta} - \theta_0}{\sigma_{\hat{\theta}}} \sim N(0,1)$$

Region de rechazo $z > z_\alpha$ cola superior, $z < -z_\alpha$ cola inferior, $|z| > z_{\frac{\alpha}{2}}$ dos colas

UNIVERSIDAD EAFIT Abierta al mundo

Intervalo de confianza con muestras pequeñas para diferencias de medias:

$$(\bar{Y}_1 - \bar{Y}_2) \pm t_{\frac{\alpha}{2}, n1+n2-2} S_P \sqrt{\frac{1}{n1} + \frac{1}{n2}}$$

Con $S_P^2 = \frac{(n1-1)S_1^2}{n1+n2-2} + \frac{(n2-1)S_2^2}{n1+n2-2}$

Intervalo de Confianza para $\frac{\sigma_2^2}{\sigma_1^2}$

$$\left[\frac{s_2^2}{s_1^2}\frac{1}{F_{n2-1,n1-1,\frac{\alpha}{2}}}, \frac{s_2^2}{s_1^2}F_{n1-1,n2-1,\frac{\alpha}{2}}\right]$$

Varianza Muestral : $s^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n}(y_i - \bar{y})^2\right]$

Intervalo de confianza con muestras grandes $\hat{\theta} \pm z_{\frac{\alpha}{2}}\sigma_{\hat{\theta}}$