UNIVERSIDAD NACIONAL ABIERTA
VICERRECTORADO ACADÉMICO
ÁREA DE MATEMÁTICA

ESTADÍSTICA APLICADA (746)
PRIMERA PARCIAL
Fecha: 11/ 10 /2008

CORRECCIÓN-MODELO DE RESPUESTAS

Objetivos 1, 2, 3 y 4.

Obj. 1 Pta. 1

Debido a que el dueño del restaurante recién abierto, El Refugio del Rey ha tenido dificultades para estimar la cantidad de comida que debe preparar cada tarde, ha decidido determinar el número medio de clientes que atiende cada noche. Seleccionó una muestra de 35 noches, que arrojaron los siguientes resultados: una media de 71 clientes con una varianza de 3,76.
¿Cuál es el intervalo que tiene 68,26 % de probabilidad de incluir la media de la población?

Solución

Un intervalo que tiene 68,26 % de probabilidad de incluir la media de la población, es equivalente a pedir la construcción de un intervalo de confianza del 68,28 %. Es claro que $1-\alpha = 0,6826$ por lo tanto $\alpha = 0{,}3174$. Usando una tabla de distribución normal tenemos que,

$$z_{\alpha/2} = 1.$$

Como el intervalo de confianza viene dado por,

$$\overline{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \overline{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

tenemos que,

$$71 - \frac{3{,}76}{\sqrt{35}} < \mu < 71 + \frac{3{,}76}{\sqrt{35}}.$$

Luego, el intervalo es $(71 \pm 0{,}6356)$.

Obj. 2 Pta. 2

José Hernández, el nuevo gerente de MacDonal's en el Valle, está interesado en el porcentaje de clientes totalmente satisfechos con el servicio. El gerente anterior tenía 86 % de clientes totalmente satisfechos, José asegura que lo mismo se cumple hoy. José obtuvo una muestra de 187 clientes y encontró que 157 estaban satisfechos por completo con el servicio. Con un nivel de confianza de 1 %. ¿Puede mantener José su afirmación?

Solución

La proporción de clientes totalmente satisfechos encontrados por José es $p = \dfrac{157}{187} = 0{,}8396$.
La fracción $1-\alpha$ se llama coeficiente de confianza o grado (nivel) de confianza. Para un nivel de confianza de 1 % el nivel de significancia es $\alpha = 0{,}99$

1) $H_0 : p = 0{,}86$ y $H_1 : p \neq 0{,}86$.
2) Para $\alpha = 0{,}99$ se tiene $z_{\alpha/2} \approx 0{,}01$.
3) Se rechaza H_0 si $z = \dfrac{\widehat{p}-p}{\sqrt{pq/n}}$ calculada cumple con $z < -z_{\alpha/2}$ ó $z > z_{\alpha/2}$
4) $z = \dfrac{0{,}8396 - 0{,}86}{\sqrt{(0{,}86)(0{,}14)/187}} = -0{,}80397.$
5) Como $z < z_{\alpha/2}$ se rechaza la hipótesis nula, por lo tanto José no puede mantener su afirmación.
¿Tiene sentido utilizar un nivel de confianza tan bajo en una prueba de hipótesis?

Obj. 3 Pta. 3

El tiempo promedio para localizar información de vuelo en los sitios de Internet de las principales aerolíneas de Venezuela es por lo común de dos a tres minutos. A continuación se muestran los resultados muestrales de los tiempos para Conviasa y Santa Barbara Airlines.

parámetros	Conviasa	Santa Barbara Airlines
$\overline{x}$	2,5 minutos	2,1 minutos
s	0,8 minutos	1,1 minutos
n	22	20

¿Existe una diferencia significativa entre los tiempos promedios para estas dos aerolíneas?

Solución

Observe que para esta prueba de hipótesis no se ha dado el nivel de significancia, α, este es asignado por el investigador antes de realizar la prueba. Cuando no se hace referencia explicita de α, se asume un valor de $\alpha = 0{,}05$ o $\alpha = 0{,}01$, valores que se usan en la práctica. Se debe utilizar una estadística t por el número de observaciones.

1) $H_0 : \mu_1 = \mu_2$ y $H_1 : \mu_1 \neq \mu_2$.
2) Para $\alpha = 0{,}05$ se tiene $t_{\alpha/2} = 2{,}021$ y para $\alpha = 0{,}01$ se tiene $t_{\alpha/2} = 2{,}704$, ambas para $22 + 20 - 2 = 40$ grados de libertad (gl).
3) Se rechaza H_0 para $\alpha = 0{,}05$ si $t < -2{,}021$ ó $t > 2{,}021$ y se rechaza H_0 para $\alpha = 0{,}01$ si $t < -2{,}704$ ó $t > 2{,}704$.
4) $t = \dfrac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$, $t = \dfrac{2,5 - 2,1}{\sqrt{\frac{0{,}8^2}{22} + \frac{1{,}1^2}{20}}} = 1{,}336374$.
5) Como $t = 1{,}336374$ calculada satisface; $-2{,}021 < t < 2{,}021$ para $\alpha = 0{,}05$ y para $\alpha = 0{,}01$ $-2{,}704 < t < 2{,}704$. No hay suficiente evidencia para rechazar la hipótesis nula, por lo tanto, se puede concluir que no hay diferencia significativa entre los tiempos promedios para acceder a la información por Internet entre las aerolíneas estudiadas, a estos niveles de significancia.

Nota: el estudiante aprueba el objetivo si responde correctamente para algún α.

Obj. 4 Pta. 4

La siguiente tabla muestra la demanda semanal de un producto,

18	20	22	27	22
25	22	27	25	24
26	23	20	24	26
27	25	19	21	25
26	25	31	29	25
25	28	26	28	24

¿Sigue la demanda de este producto una distribución normal?

Solución

Una forma para determinar si la variable aleatoria sigue una distribución normal consiste en realizar una prueba de bondad de ajuste, que implica el uso de la distribución χ^2 para hacer el contraste. No se hace referencia al nivel de significación por lo tanto podemos elegir entre $\alpha = 0{,}05$ ó $\alpha = 0{,}01$, valores usados en la práctica. Como estamos ajustando a una distribución continua, debemos agrupar los datos en categorías (Estadística General 746).

Usamos la fórmula de Sturges para tener una idea del tamaño apropiado de los intervalos de clases,

$$c = \frac{rango}{1 + 3{,}322 \log N} = \frac{31 - 18}{1 + 3{,}322 \log 30} = 2{,}2007,$$

el tamaño de clase apropiado es dos.

La siguiente tabla resume la data y los valores necesarios para estimar los parámetros poblacionales $\overline{x}$ y s^2,

intervalo	f_i	x_i	$f_i x_i$	$(x_i - \overline{x})^2$
$18 < 20$	2	19	38	36
$20 < 22$	3	21	63	16
$22 < 24$	4	23	92	4
$24 < 26$	10	25	250	0
$26 < 28$	7	27	189	4
$28 < 30$	3	29	87	16
$30 < 32$	1	31	31	36
Total	30		750	112

Así, $\overline{x} = \dfrac{\sum f_i x_i}{n} = 25$, $s^2 = \dfrac{\sum (x_i - \overline{x})^2}{n-1} = 3{,}862$, y $s = 1{,}97$.

Procedemos ahora a determinar las observaciones esperadas, determinando primero las probabilidades correspondientes a cada intervalo.

$P_1 = P(x < 20) = P(z < (20 - 25)/1{,}97) = P(z < -2{,}54) = 0{,}0055$
$P_2 = P(20 < x < 22) = P((20 - 25)/1{,}97 < z < (22 - 25)/1{,}97) = P(-2{,}54 < z < -1{,}52) = 0{,}0588$
$P_3 = P(22 < x < 24) = P((22 - 25)/1{,}97 < z < (24 - 25)/1{,}97) = P(-1{,}52 < z < -0{,}51) = 0{,}2407$
$P_4 = P(24 < x < 26) = P((24 - 25)/1{,}97 < z < (26 - 25)/1{,}97) = P(-0{,}51 < z < 0{,}51) = 0{,}3900$
$P_5 = P(26 < x < 28) = P((26 - 25)/1{,}97 < z < (28 - 25)/1{,}97) = P(0{,}51 < z < 1{,}52) = 0{,}2407$
$P_6 = P(28 < x < 30) = P((28 - 25)/1{,}97 < z < (30 - 25)/1{,}97) = P(1{,}52 < z < 2{,}54) = 0{,}0588$
$P_7 = P(x > 30) = P(z > (30 - 25)/1{,}97) = P(z > 2{,}54) = 0{,}0055$

Con estos valores de probabilidades construimos la siguiente tabla,

intervalo	O_i	P_i	E_i
$18 < 20$	2	0,0055	0.2
$20 < 22$	3	0,0588	1,8
$22 < 24$	4	0.2407	7.2
$24 < 26$	10	0.3900	11.7
$26 < 28$	7	0.2407	7.2
$28 < 30$	3	0.0588	1.8
$30 < 32$	1	0.0055	0.2

Según estos resultados parciales debemos combinar algunos de estos resultados, para que ninguna frecuencia esperada sea menor que cinco, por lo tanto el número de intervalos se reduce de siete a tres

intervalo	O_i	E_i
$18 < 24$	9	9,2
$24 < 26$	10	11.7
$26 < 32$	11	9.2

Ahora podemos plantear nuestra estructura para contraste de hipótesis,

1) H_0 :"la variable aleatoria sigue una distribución normal" y H_1 :"la variable aleatoria no sigue una distribución normal".
2) Para $\alpha = 0{,}05$ se tiene $\chi^2_\alpha = 5{,}99$ y para $\alpha = 0{,}01$ se tiene $\chi^2_\alpha = 9{,}21$, ambas para $3 - 1 = 2$ gl.
3) Se rechaza H_0 para $\alpha = 0{,}05$ si $\chi^2 > 5{,}99$ y se rechaza H_0 para $\alpha = 0{,}01$ si $\chi^2 > 9{,}21$.
4) $\chi^2 = \dfrac{(9-9{,}2)^2}{9{,}2} + \dfrac{(10-11{,}7)^2}{11{,}7} + \dfrac{(11-9{,}2)^2}{9{,}2} = 0{,}6035$
5) Como $\chi^2 < \chi^2_\alpha$ para cualquier valor de α elegido, no hay suficiente evidencia para rechazar la hipótesis nula, por lo tanto, se puede concluir que la variable aleatoria sigue una distribución normal, al nivel de significancia α elegida.

Nota: el estudiante aprueba el objetivo si responde correctamente para algún α.

FIN DEL MODELO