Nombre: *Javier Rivilla Arredondo*
DNI: *53247378D*
Email: *jra48@alu.ua.es*

PRÁCTICA 6

<u>Ejercicio 6.1</u>: El control de calidad de una fábrica de pilas y baterías sospecha que hubo defectos en la producción de un modelo de batería para teléfonos móviles, bajando su tiempo de duración. Hasta ahora el tiempo de duración en conversación seguía una distribución normal con media 300 minutos y desviación típica 30 minutos. Sin embargo, en la inspección del último lote producido, antes de enviarlo al mercado, se obtuvo que de una muestra de 60 baterías el tiempo medio de duración en conversación fue de 290 minutos. Suponiendo que ese tiempo sigue siendo Normal con la misma desviación típica, contrasta a un nivel de significación del 2% si las sospechas del control de calidad son ciertas.

Solución:

Realizado en clase de teoría

<u>Ejercicio 6.2</u>: En municipios de 2000 a 5000 habitantes de todo el territorio nacional se calcula que la media de horas vistas de televisión por hogar al día es de 4.31 horas. Sea una zona de la que se tomó una muestra de 120 municipios de aquellas dimensiones, con un promedio de 3.86 horas de televisión al día y con desviación típica de 1.02. Contrasta con un nivel de significación del 1 por ciento si en dicha zona se puede considerar que el promedio es igual al de todo el territorio.

Solución:

H0 : u = 4.31
H1 : u != 4.31
Desviación conocida = 1.02

Sabemos que la desviación típica es conocida por lo que tenemos que utilizar la siguiente fórmula de **Estadístico de prueba:**

$H_0 : \mu = \mu_0$ σ^2 conocida	$Z_0 = \frac{\overline{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}}$	$H_1 : \mu \neq \mu_0$	$\lvert Z_0\rvert > Z_{\frac{\alpha}{2}}$	$2P(Z \geq \lvert Z_0\rvert)$
		$H_1 : \mu > \mu_0$	$Z_0 > Z_\alpha$	$P(Z \geq Z_0)$
		$H_1 : \mu < \mu_0$	$Z_0 < -Z_\alpha$	$P(Z \leq Z_0)$

α = 0.01
Z_0 = (3.86 - 4.31) / (1.02 / raiz (120)) = **-4.8328**
$Z_{0.01/2}$ = IDF.Normal(0.0995, 0, 1) = **2.57**

Criterio de rechazo, si se cumple $|Z_0| > Z_{\alpha/2}$ se rechaza el criterio, como antes hemos calculado que $|Z_0|$ = **4.8328** podemos decir que 4.8328 > 2.57 por lo que se cumple, así que se rechaza el criterio de rechazo.

Para verificarlo vamos a comprobarlo con el P-valor, de esta forma el P-valor debería de darnos menos que el α:

2 * P(Z >= |-4.8328|) = 2 * (1 - P(Z <= 4.83)) = 2 * (1 - 0.999999) = **0.00002**

Como podemos comprobar α > P-valor → 0.01 > 0.00002 por lo que se rechaza la hipótesis nula.

Ejercicio 6.3: En una universidad están pensando aumentar el número de terminales conectadas a internet. La red actual estaba pensada para una media de conexiones en las horas punta de 1000. En una serie de días lectivos consecutivos y a la misma hora se ha observado el número de terminales de una universidad conectados a Internet. Los resultados han sido los siguientes: 1027 1023 1369 950 1436 957 834 821 882 942 904 984 1067 570 1063 1307 1212 1045 1047 1178 633 501 565 1039 1000 1227 1118 843 696 820 1092 834 Si suponemos normalidad explica qué debe hacer la universidad.

H0 : u = 1000
H1 : u > 1000

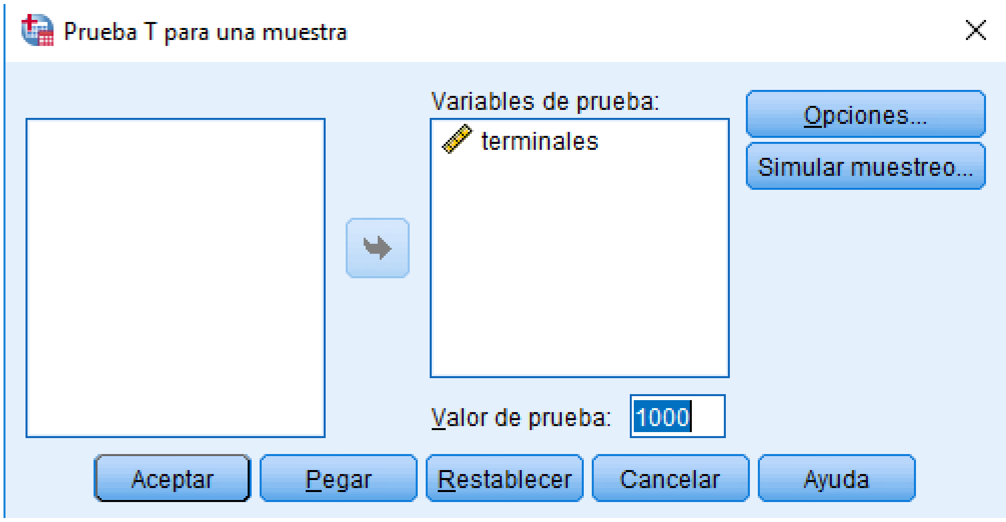
Solución:

	terminales
1	1027
2	1023
3	1369
4	950
5	1436
6	957
7	834
8	821
9	882
10	942
11	904
12	984
13	1067
14	570
15	1063
16	1307
17	1212
18	1045
19	1047
20	1178
21	633
22	501
23	565
24	1039
25	1000
26	1227
27	1118
28	843
29	696
30	820
31	1092
32	834

Como desconocemos la desviación:

$H_0 : \mu = \mu_0$ σ^2 desconocida	$t_0 = \frac{\overline{X}-\mu_0}{\frac{S}{\sqrt{n}}}$	$H_1 : \mu \neq \mu_0$	$\lvert t_0\rvert > t_{\frac{\alpha}{2},n-1}$	$2P(t_{n-1} \geq \lvert t_0\rvert)$
		$H_1 : \mu > \mu_0$	$t_0 > t_{\alpha,n-1}$	$P(t_{n-1} \geq t_0)$
		$H_1 : \mu < \mu_0$	$t_0 < -t_{\alpha,n-1}$	$P(t_{n-1} \leq t_0)$

Añadimos los datos al SPSS → **Comparar medias** → **Prueba T para una muestra**:



En valor de prueba añadimos el promedio = 1000

Prueba para una muestra

	Valor de prueba = 1000					
					95% de intervalo de confianza de la diferencia	
	t	gl	Sig. (bilateral)	Diferencia de medias	Inferior	Superior
terminales	-,794	31	,433	-31,688	-113,08	49,71

El intervalo del 95% es de {**-113.08, 49.71**}.

De los resultados obtenidos con el SPSS tenemos entonces que 2P(T ≥ |-0.794|)=0.4333
Sin embargo nosotros tenemos el siguiente contraste unilateral:

H0 : u = 1000
H1 : u > 1000

$P(t_{n-1} >= t_0) = P(t_{31} >= -0.794) = 1 - P(t_{31} <= -0.794) = 1 - CDF.T(-0.794, 31) = 1 - 0.2166 =$ **0.7934**

Dado que α = 1 - 0.95 = 0.05 obtenemos que P-valor > α → 0.7934 > 0.05 por lo que se acepta la hipótesis nula de aumentar los terminales en la universidad.

<u>Ejercicio 6.4</u>: Un centro de investigación ha diseñado un programa de fisioterapia con la Wii, para que los pacientes de rehabilitación puedan hacer los ejercicios en casa. Se sabe que, para cierto tipo de dolencia, la rehabilitación clásica dura una media de 4 meses. El centro desea saber si con el uso de dicho programa el tiempo de recuperación para ese tipo de dolencia es menor que si la rehabilitación se hace de forma clásica. Para ello utilizaron el programa sobre 2500 pacientes y se obtuvo que la duración media de rehabilitación fue de 3.5 meses con una desviación típica de 0.7 meses. Realizando el contraste de hipótesis adecuado explica si los datos obtenidos con el uso de la Wii mejoran los datos estimados previamente.

Solución:

Contraste de hipótesis:

H_0 : u = 4
H_1 : u < 4

$H_0 : \mu = \mu_0$	$Z_0 = \frac{\overline{X}-\mu_0}{\frac{S}{\sqrt{n}}}$	$H_1 : \mu \neq \mu_0$	$\lvert Z_0 \rvert > Z_{\frac{\alpha}{2}}$	$2P(Z \geq \lvert Z_0 \rvert)$
		$H_1 : \mu > \mu_0$	$Z_0 > Z_\alpha$	$P(Z \geq Z_0)$
		$H_1 : \mu < \mu_0$	$Z_0 < -Z_\alpha$	$P(Z \leq Z_0)$

u_0 = 4
X = 3.5
S = 0.7
n = 2500

Z_0 = (3.5 - 4) / [0.7/raiz(2500)] = **-35.714**

Criterio de rechazo → Z_0 < -Zα suponemos que α = 0.05 por lo que Z_0 = **-35.714** y $Z_{0.05}$ = -1.645

Como -Zα > Z_0 se rechaza el criterio de rechazo.

Calculamos el P-valor → P(Z <= Z_0) = P(Z <= -35.71) = CDF.Normal(-35.71, 0, 1) = **0**
Como α > P-valor se rechaza la hipótesis nula de los datos estimados previamente de la WII.

Ejercicio 6.5: En un servicio médico se sabe que el 15 por ciento de las visitas que requieren la realización de una radiografía llevan consigo un ingreso hospitalario. Un centro de investigación ha diseñado un sistema experto de ayuda al diagnóstico para interpretar radiografías. El servicio médico durante el año pasado puso en marcha dicho sistema como ayuda previa al diagnóstico realizado por el especialista. De 2450 visitas, el sistema indicó que 350 requerían ingreso hospitalario. Contrasta si los datos obtenidos con el sistema experto están en concordancia con la cifra inicial conocida por el servicio médico.

Solución:

Como estamos hablando de proporciones de muestras grandes:

H_0 : p = 15%
H_1 : p != 15%

Usaremos el siguiente contraste de hipótesis:

$H_0 : p = p_0$	$Z_0 = \frac{p^* - p_0}{\sqrt{\frac{(1-p_0)p_0}{n}}}$	$H_1 : p \neq p_0$	$\lvert Z_0 \rvert > Z_{\frac{\alpha}{2}}$	$2P(Z \geq \lvert Z_0 \rvert)$
		$H_1 : p > p_0$	$Z_0 > Z_\alpha$	$P(Z \geq Z_0)$
		$H_1 : p < p_0$	$Z_0 < -Z_\alpha$	$P(Z \leq Z_0)$

p_0 = 15%
p* = 350/2450 = 1/7
n = 2450

Z_0 = (1/7 - 0.15) / (Raiz(((1 - p_0) * p_0) / n) = -7.14x10^-3 / 7.21x10^-3 = **-0.990291**

Como no nos dan el valor de significación suponemos que **α = 0.05**

Calculamos el P-valor:
2P(Z >= |Z_0|) = 2 * (1 - P(Z <= |Z0|) = 2 * (1 - P(Z <= **0.990291**)) = 2 * (1 - 0.8389) = 2 * 0.1611 = **0.3222.**

Si α > P-valor se rechaza la hipótesis nula, como en este caso, α < 0.3222 por lo que, se acepta la hipótesis nula, concordando los datos obtenidos con la cifra inicial conocida por el servicio médico.

Ejercicio 6.6: En un centro de investigación se está trabajando en el desarrollo de una aplicación de visión robótica para la automatización de vehículos. Para la detección de obstáculos se están utilizando dos métodos distintos (método 1 y método 2). Con el fin de analizar si existen diferencias significativas entre los resultados obtenidos con ambos métodos

se implementaron en un vehículo y se probaron en circuitos cerrados. La tasa de error de dichos métodos, en cada caso, expresada en tanto por ciento, fue:
Método 1: 12, 14, 10, 25, 11, 8, 17, 11.
Método 2: 10, 9, 15, 12, 15, 7, 27, 11. No se han podido determinar las desviaciones típicas poblacionales pero se considera que son iguales. Estudia planteando el correspondiente contraste de hipótesis y de forma razonada qué conclusiones se pueden obtener a un nivel de significación del 1 por ciento si se supone normalidad.

Solución:

El contraste que he supuesto es el siguiente:

Sea u_1 = método 1 aplicado para la detección de obstáculos
Sea u_2 = método 2 aplicado para la detección de obstáculos

$H_0 : u_1 = u_2$
$H_1 : u_1 \text{ != } u_2$

$\alpha = 0.01$, $n_1 = n_2 = 8$, supuesto de normalidad y se desconocen desviaciones típicas poblaciones pero se sabe que son iguales:

TABLA 2 Bloque 2

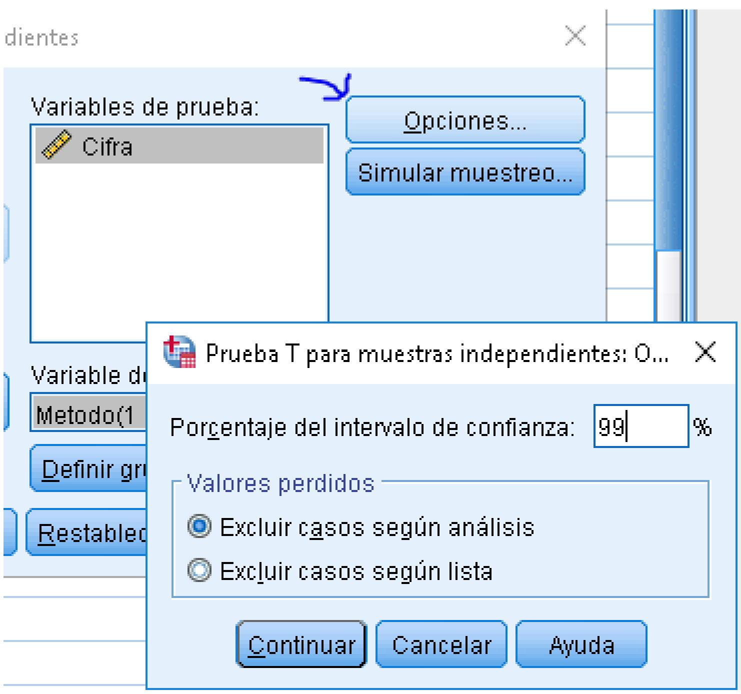
$H_0 : \mu_1 = \mu_2$ $\sigma_1^2 = \sigma_2^2$ desconocidas	$t_0 = \frac{\overline{X}_1 - \overline{X}_2}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$	$H_1 : \mu_1 \neq \mu_2$	$\lvert t_0 \rvert > t_{\frac{\alpha}{2}, n_1+n_2-2}$	$2P(t_{n_1+n_2-2} \geq \lvert t_0 \rvert)$
		$H_1 : \mu_1 > \mu_2$	$t_0 > t_{\alpha, n_1+n_2-2}$	$P(t_{n_1+n_2-2} \geq t_0)$
		$H_1 : \mu_1 < \mu_2$	$t_0 < -t_{\alpha, n_1+n_2-2}$	$P(t_{n_1+n_2-2} \leq t_0)$

Como desconocemos la desviación podemos usar el SPSS, primero añadimos los datos:

	Cifra	Metodo
1	12	1
2	14	1
3	10	1
4	25	1
5	11	1
6	8	1
7	17	1
8	11	1
9	10	2
10	9	2
11	15	2
12	12	2
13	15	2
14	7	2
15	27	2
16	11	2

Donde en cada cifra la corresponde a cada método.

Analizar → Comparar medias → Prueba T para muestras independientes



El valor de significación es de 1%, es decir, $\alpha = 0.01 \rightarrow (1-\alpha) * 100 = 99\%$ valor de confianza.

Prueba de muestras independientes

		Prueba de Levene de igualdad de varianzas		prueba t para la igualdad de medias						
									99% de intervalo de confianza de la diferencia	
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Diferencia de error estándar	Inferior	Superior
Cifra	Se asumen varianzas iguales	,053	,821	,086	14	,933	,250	2,902	-8,388	8,888
	No se asumen varianzas iguales			,086	13,719	,933	,250	2,902	-8,415	8,915

Como se asumen varianzas iguales nos quedamos con "Se asumen varianzas iguales":

prueba t para la igualdad de medias						
					99% de intervalo de confianza de la diferencia	
t	gl	Sig. (bilateral)	Diferencia de medias	Diferencia de error estándar	Inferior	Superior
,086	14	,933	,250	2,902	-8,388	8,888

t_0

Como observamos los valores dados son los siguientes:

$T_0 = 0.086$
Gl(grados de libertad) = $n_1+n_2-2 = 8+8-2 = 14$
Sig. (Bilateral) = P-valor contraste bilateral

Según los datos obtenidos para varianzas iguales por el SPSS 2P(T >= |0.086|) = 0.933 donde T es una t de Student con 14 grados de libertad. Pero en este caso el P-valor del contraste bilateral planteado coincide con el SPSS por lo que nuestro P-valor será **0.933.**

Si α > P-valor se rechaza la hipótesis nula, en este caso α = 0.01 y p-Valor = 0.933, por lo que α < P-valor se acepta la hipótesis nula.

Ejercicio 6.7: El equipo ING Renault F1, ha investigado para esta temporada un nuevo método para la refrigeración del motor del R-28. Se decidió no aceptar este nuevo método a no ser que su rendimiento medio fuera superior al del actual sistema. Se probó este nuevo método, manteniendo todo constante, y se obtuvieron los siguientes resultados:
Método actual: 232, 266, 244, 235, 226, 257, 255.
Método nuevo: 210, 277, 262, 279, 223, 224, 281.
Estudia qué conclusiones se pueden obtener al nivel de significación del 2 por ciento, si se supone normalidad y desviaciones típicas distintas.

Solución:

Contraste supuesto:

Sea u_1 = método actual para la refrigeración del motor
Sea u_2 = método nuevo para la refrigeración del motor

$H_0 : u_1 = u_2$
$H_1 : u_1 < u_2$

α = 0.02, n1 = n2 = 7, supuesto de normalidad y se desconocen las desviaciones típicas poblaciones pero se sabe que son distintas:

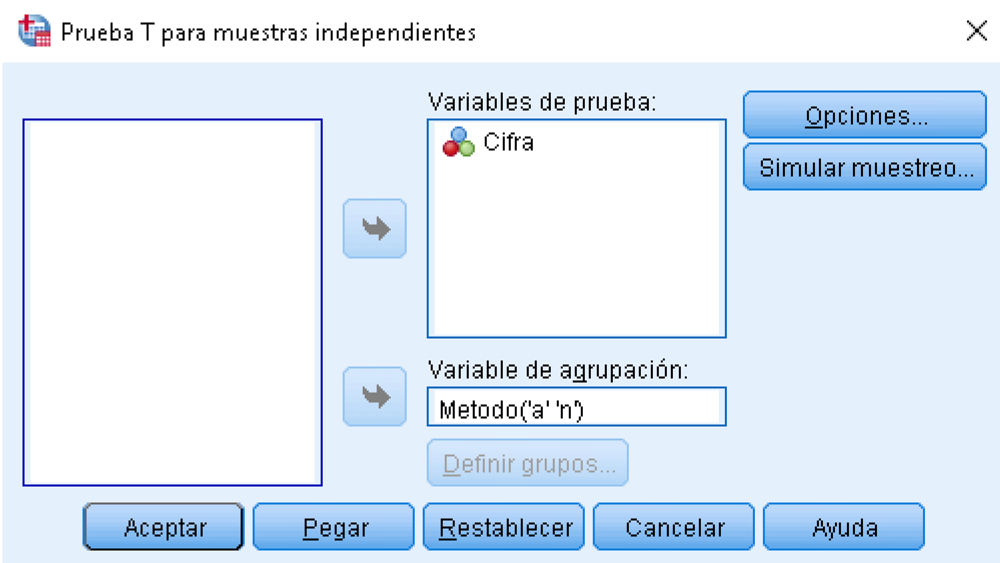
$H_0 : \mu_1 = \mu_2$ $\sigma_1^2 \neq \sigma_2^2$ desconocidas	$t_0 = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$	$H_1 : \mu_1 \neq \mu_2$	$\lvert t_0 \rvert > t_{\frac{\alpha}{2},\nu}$	$2P(t_\nu \geq \lvert t_0 \rvert)$
		$H_1 : \mu_1 > \mu_2$	$t_0 > t_{\alpha,\nu}$	$P(t_\nu \geq t_0)$
		$H_1 : \mu_1 < \mu_2$	$t_0 < -t_{\alpha,\nu}$	$P(t_\nu \leq t_0)$

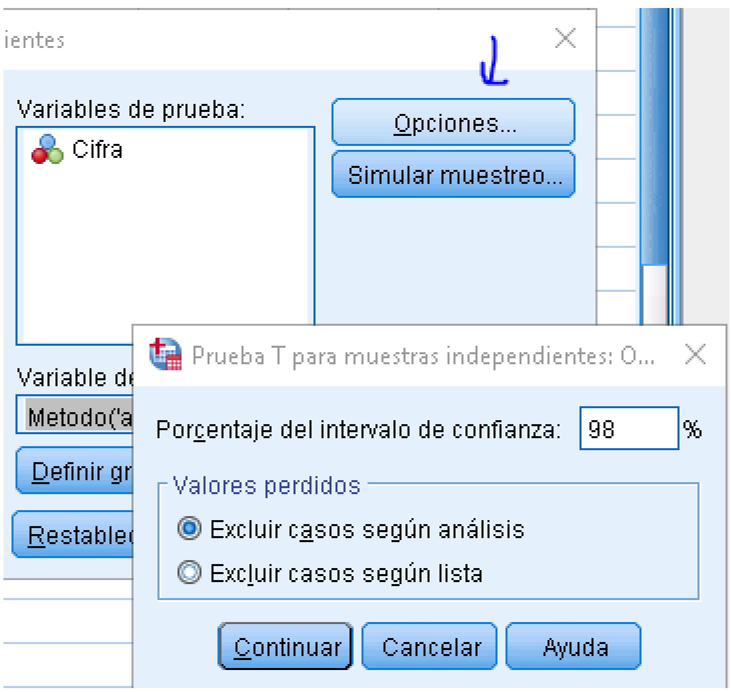
Como no conocemos las desviación típica usaremos el SPSS, primero añadimos los datos:

	Cifra	Metodo
1	232	a
2	266	a
3	244	a
4	235	a
5	226	a
6	257	a
7	255	a
8	210	n
9	277	n
10	262	n
11	279	n
12	223	n
13	224	n
14	281	n

Donde a se refiere al método actual u n se refiere al método nuevo.

Analizar → Comparar medias → Prueba T para muestras independientes





Prueba de muestras independientes

		Prueba de Levene de igualdad de varianzas		prueba t para la igualdad de medias						
									98% de intervalo de confianza de la diferencia	
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Diferencia de error estándar	Inferior	Superior
Cifra	Se asumen varianzas iguales	13,254	,003	-,454	12	,658	-5,857	12,902	-40,446	28,732
	No se asumen varianzas iguales			-,454	8,641	,661	-5,857	12,902	-42,575	30,861

Como se asumen desviación típicas distintas obtenemos el siguiente:

					prueba t para la igualdad de medias	
					98% de intervalo de confianza de la diferencia	
t	gl	Sig. (bilateral)	Diferencia de medias	Diferencia de error estándar	Inferior	Superior
-,454	12	,658	-5,857	12,902	-40,446	28,732
-,454	8,641	,661	-5,857	12,902	-42,575	30,861

$t_0 = -0.454$
$v = 8.641$
Sig. (bilateral) → P-valor contraste bilateral $H_1 : u_1 != u_2$

Como bien sabemos el SPSS nos da un contraste bilateral, es decir, el P-valor = 2P(T >= $|t_0|$) = 0.661. Pero nosotros tenemos el contraste unilateral $H_1 : u_1 < u_2$,es decir, P(T <= t_0) tenemos que calcular, para ello P($T_{8.641}$ <= -0.454), teniendo en cuenta que T es simétrica respecto al cero: P-valor = P($T_{8.641}$ >= 0.454) = 0.661 / 2 = **0.3305**

El método nuevo se aceptará si supera el método actual, es decir, si α > p-Valor se rechazará la hipótesis nula, como α = 0.02:

Cómo 0.02 < 0.3305 se aceptará la hipótesis nula esto quiere decir, que el método nuevo no supera en rendimiento al método actual.

Ejercicio 6.8: Una serie de encuestas sucesivas de opinión sobre el uso de las nuevas tecnologías en la empresa habían revelado entre otros datos, el siguiente porcentaje de respuestas afirmativas a determinada pregunta. En enero de 2010 sobre una muestra de 420 empresas se obtenía un 30 por ciento de respuestas afirmativas, mientras que en enero de 2011 y en una muestra de 220 empresas se obtenía un 40 por ciento de respuestas afirmativas. Explica hasta qué punto podría aceptarse que ha aumentado significativamente la proporción de respuestas afirmativas.

Solución:

Como estamos trabajando con muestras grandes, utilizaremos las correspondientes a la tabla 3.

$H_0 : p_1 = p_2$	$Z_0 = \frac{p_1^* - p_2^*}{\sqrt{p^*(1-p^*)}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$	$H_1 : p_1 \neq p_2$	$\|Z_0\| > Z_{\frac{\alpha}{2}}$	$2P(Z \geq \|Z_0\|)$
		$H_1 : p_1 > p_2$	$Z_0 > Z_\alpha$	$P(Z \geq Z_0)$
		$H_1 : p_1 < p_2$	$Z_0 < -Z_\alpha$	$P(Z \leq Z_0)$

P1 = proporcion de enero 2010
P2 = proporcion de enero 2011

Para que aumente se tiene que cumplir que la proporcion de 2011 sea más grande que la de 2010:

$H_0 : p_1 = p2$
$H_1 : p_1 < p2$

p* = X / n = 214 / 640 = **0.3343**
p1* = 126/420 = **0.30**
p2* = 88/220 = **0.40**

n_1 = **420**
n_2 = **220**
n = 420 + 220 = **640**

Z0 = (0.30 - 0.40) / [raiz(0.3343 * (1 - 0.3343)) * raiz(1/420 + 1/220)] = -0.1 / (0.4717 * 0.08322) = **-2.5474**

Una vez esto calculamos el P-valor para comprobar la hipótesis:

$P(Z <= Z_0)$ = P(Z <= -2.5474) = 1 – P(Z <= 2.5474) = 1 – 0.9945 = 5.5*10^-3

Si α > p-valor se rechaza la hipótesis nula, por lo que como α > P-valor → 0.05 > 5.5*10^-3 se rechaza la hipótesis nula, aceptando así que ha aumentado significativamente la proporción de respuestas afirmativas.

Ejercicio 6.9: En un periódico se afirmó que entre los menores de 10 a 15 años hay más niñas usuarias de Internet que niños usuarios. Para contrastar si ese dato era cierto una importante empresa informática tomó una muestra aleatoria de 10000 menores de dichas edades y resultó que de las 6000 niñas muestreadas, 4704 habían usado Internet en los últimos tres meses, mientras que de los 4000 niños muestreados, se habían conectado a Internet 3012. Mediante el contraste de hipótesis adecuado, explica si se podría dar por cierta la afirmación del periódico.

Solución:
Realizado en clase de teoría

Ejercicio 6.10: Una empresa de marketing e Internet ha prestado sus servicios a una gran multinacional con el fin de conseguir un mejor posicionamiento de su sitio Web y aumentar el número de páginas vistas. Entre los aspectos tratados se analizó la tasa de rebotes. El estudio previo mostró que dicho sitio Web tenía hasta el momento una tasa de rebote del 32 por ciento. Después del asesoramiento técnico de la empresa se analizó la tasa de rebote tomando una muestra de 5000 visitas obteniendo que de dichas visitas 1500 vieron una única página. Mediante el contraste de hipótesis adecuado, explica si se podría dar por cierta la afirmación de que la tasa de rebote ha disminuido después de dicho asesoramiento técnico.

Solución:
Realizado en clase de teoría