EJERCICIOS PROPUESTOS

1. Suponga que se estudia la compra de una nueva maquina para una empresa. Se comprara la maquina si la proporción de la producción que necesita ser reprocesados por tener defectos es inferior al 3 %. Se examina una muestra de 50 artículos construidos por la maquina y 5 necesitan ser reprocesados . ¿ Que decisión se toma? (Se compra o no la maquina?)

2. Su ponga que una empresa desarrolla un curso de entrenamiento para sus empleados, formando dos grupos y aplicándoles dos métodos distintos de entrenamiento. El primer grupo lo componen 36 empleados que obtuvieron un puntaje promedio de 6 (en escala de 0 a 10 puntos) y una desviación estándar de 4 puntos y el segundo grupo de 40 empleados cuyo puntaje promedio fue de 8.2 y una desviación de 4.3. Se puede afirmar que el método aplicado al segundo grupo es superior al aplicado al primero? Que supuestos debe de tener en cuenta?

3. Los ingenieros de una ensambladora de automóviles requieren decidir sobre cuál de dos de las marcas de neumáticos deben comprar. La marca FB o la marca KT. Con el fin de tomar una decisión basada en evidencias estadísticas, deciden realizar un experimento en el que usan 12 neumáticos de cada marca. Los neumáticos se utilizan hasta su terminación. Los resultados obtenidos son los siguientes:

   Marca FB: 40.5 38.4 44.0 34.9 44.0 44.7 44.0 47.1 39.8 43.9 44.2 40.2
   Marca KT: 41.8 41.6 31.5 48.7 40.8 31.2 36.5 36.2 32.8 36.3 38.6 30.5

   Cuál marca de neumáticos recomendaría comprar. Justifique su respuesta. Suponga que la distancia recorrida por un neumático se distribuye aproximadamente normal y un $\alpha = 0{,}05$.

4. Un ingeniero desea establecer si existen diferencias entre dos métodos diferentes de realizar el ensamble de una casa prefabricada. Para comprobarlo recoge información de ambos métodos que se presentan a continuación:

   Procedimiento estándar: 35, 31, 29, 25, 34, 40, 27, 32, 31.
   Nuevo procedimiento: 32, 37, 35, 28, 41, 44, 35, 31, 34.

   Presentan los datos suficiente evidencia estadística para afirmar que el nuevo método es más eficiente que el estándar? (utilice un $\alpha = 0{,}05$).

5. Un director de un gimnasio quiere determinar si un instructor de ejercicio debe ser contratado o no para su campaña estrella "Reducción de peso", Para tomar la decisión le dice que pruebe con 16 de las personas que habitualmente concurren tomadas al azar. Los datos que se tomaron antes (x_1) y después (x_2) de haber realizado un mes de ejercicios son los siguientes:

id	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
x_1	104	89	84	106	90	96	79	90	85	76	91	82	100	89	121	72
x_2	98	85	85	103	88	95	79	90	82	76	89	81	99	86	111	70

   Emplee y realice las pruebas de hipótesis a un nivel de significancia del 0.01 para determinar si el programa que ofrece el nuevo instructor es eficaz. Suponga que la variable peso se distribuye aproximadamente normal.

.

6. Se realizan pruebas de un nuevo lector láser manual para uso en inventarios y el lector utilizado actualmente, con el fin de decidir si se adquiere el primero. Se obtienen los datos siguientes sobre el número de códigos de barra de 7 pulgadas que pueden leerse por segundo. Sea X_1: número de códigos leído por segundo con el dispositivo nuevo y X_2 el correspondiente al dispositivo antiguo.

$$n_1 = 61\ ; \qquad \bar{x}_1 = 40\ ; \qquad s_1^2 = 24{,}9$$
$$n_2 = 61\ ; \qquad \bar{x}_2 = 32\ ; \qquad s_2^2 = 22{,}7$$

De acuerdo con la información suministrada, es posible preferir alguno de ellos?. En caso de poderlo realizar con cual se quedaría? Justifique su respuesta. En cada caso determine las pruebas de hipótesis, el estadístico de prueba apropiado, el $valor - p$ obtenido y las conclusiones resultantes.

7. Un empresario registro el número de artículos producidos durante 10 días, para un grupo de 15 obreros que trabajaban con base en un salario fijo (Grupo 1). El industrial introdujo un plan de incentivos para otros 15 obreros y registro su producción durante otros 10 días (Grupo 2). El número de artículos producidos por cada uno de los grupos fue :

G1	75	76	74	80	72	78	76	73	72	75
G2	86	78	86	84	81	79	78	84	88	80

Suponiendo que los salarios pagados a cada grupo son equivalentes. Se puede concluir que el plan de incentivos es efectivo?

8. En una muestra de 200 clientes, el 20 % indica una preferencia por tamaño especial de pizza. Con posterioridad a una campaña publicitaria realizada en radio y televisión promoviendo dicho producto, se selecciono una muestra de igual tamaño. En esta ultima muestra el 22 % de los clientes indico preferencia por el producto. De acuerdo con estos resultados y un nivel de significancia del 5 % , podría decirse que la campaña publicitaria no fue efectiva?

9. Los siguientes son los datos de las horas hombre que se pierden en promedio por accidentes en 10 plantas industriales antes (Ant) y después (Des) de la implantación de un programa de seguridad industrial:

id	1	2	3	4	5	6	7	8	9	10
Ant	45	73	46	124	30	57	83	34	26	17
Des	37	62	44	119	36	55	77	19	24	11

Utilice un nivel de significancia de 0.05 para probar si el programa de seguridad implantado es eficaz. Suponga que esta variable se distribuye aproximadamente normal.

.

10. La compañía de dulces Mars publica en su sitio web información relacionada con los porcentajes de los distintos colores de sus dulces M&M para la variedad de chocolate con leche.

Color contenido en la bolsa	café	amarillo	rojo	azul	naranja	verde
Porcentaje (%)	13	14	13	24	20	16

Se realiza una verificación mediante el conteo delos dulces contenidos e n una bolsa de 14 onzas de dulces M&M, obteniendo los siguientes resultados: 70 duces cafés, 72 amarillos, 61 rojos, 118 azules, 108 naranjas y 85 verdes.

Se podria afirmar que los datos anteriores respaldan la información suministrada por la compañía en su sitio web? Sustente su respuesta.

11. En una línea de producción los artículos se inspeccionan en forma periódica con el fin de detectar defectos. La siguiente secuencia de artículos defectuosos (D) y no defectuosos (N) corresponde a la producción de uno de los turnos.

```
D D N N N N D N N D D N N N N N D D D D N N D N N N N D N D N N N N N N D N N N D D N N
N N N N D N D N N N N D D D D D N D D N N N N N N N D D D D D D D D D N N N N N N D D N
```

Se puede afirmar que los datos no presentan patrón alguno y que la generación de artículos defectuosos se debe al azar? . Utilice un $\alpha = 0{,}05$.

12. En una planta ensambladora de camiones la supervisión diaria de las soldaduras generó la siguiente información :

	Número de soldaduras		
Turno	Alta calidad	Moderada calidad	Baja calidad
dia	470	191	42
tarde	445	171	28
noche	257	129	17

¿Se puede concluir que la calidad varia con los turnos?, en otras palabras se puede concluir que la calidad de las soldaduras es independiente de los turnos? . Utilice un nivel de significancia $\alpha = 0{,}05$.

13. Los siguientes datos corresponde a las notas obtenidas por un grupo de estudiantes de la asignatura Matemáticas Fundamentales. Si la distribución de los datos es normal, podría afirmar que la prueba realizada es una prueba normalizada. En caso contrario serviría para estudiar problemas relacionados con su aprendizaje. Para un $\alpha = 0; 05$, se podría afirmar que los datos proceden de una distribución normal? . Si se requiere realizar una prueba de hipótesis sobre la media de la nota $H_o : \mu \leq 3{,}3$ vs $H_a : \mu > 3{,}3$, ¿Que prueba se realizaría?

```
3.4, 2.8, 4.2, 2.1, 2.8, 2.4, 3.5, 4.2, 3.1, 4.1, 2.4, 3.4, 4.1, 4.0, 2.4, 4.1, 3.4, 4.4,
3.8, 3.7, 2.2, 3.6, 2.3, 3.7, 2.8, 4.1, 2.3, 4.6, 4.6, 5.2, 2.4, 2.4, 2.7, 3.8, 4.6, 4.4,
4.2, 4.4, 2.4, 3.3, 3.8, 2.9, 3.1, 2.7, 3.6, 3.8, 4.4, 3.9, 2.8, 3.7
```

FORMULARIO PRUEBAS DE HIPÓTESIS PARAMÉTRICAS

$$Z_o = \frac{\bar{X} - \mu_o}{\sigma/\sqrt{n}} \sim N(0,1) \tag{1}$$

$$Z_o = \frac{\bar{X} - \mu_o}{s/\sqrt{n}} \sim N(0,1) \tag{2}$$

$$T_o = \frac{\bar{X} - \mu_o}{s/\sqrt{n}} \sim t_{v=n-1} \tag{3}$$

$$X_o^2 = \frac{(n-1)S^2}{\sigma_o^2} \sim \chi^2_{v=n-1} \tag{4}$$

$$Z_o = \frac{X - n\,p_o}{\sqrt{n\,p_o}} \sim N(0,1) \tag{5}$$

$$Z_o = \frac{\widehat{p} - p_0}{\sqrt{p_o(1-p_o)/n}} \sim N(0,1)$$

$$Z_o = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta_o}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \sim N(0,1) \tag{6}$$

$$T_o = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta_o}{s_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \sim t_{v=n_1+n_2-2} \tag{7}$$

donde: $s_p^2 = \dfrac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}$

$$T_o = \frac{(\bar{x}_1 - \bar{x}_2 - \Delta_o}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} \sim t_{v^*} \tag{8}$$

donde: $v^* = \dfrac{(s_1^2/n_1 + s_2^2/n_2)^2}{\dfrac{(s_1^2/n_1)^2}{n_1-1} + \dfrac{(s_2^2/n_2)^2}{n_2-1}}$

$$T_o = \frac{\bar{d} - \Delta_o}{s_d^2} \sim t_{v=n-1} \tag{9}$$

donde: $\bar{d} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} d_i$ y $d_i = x_1 - x_2$

$$Z_o = \frac{(\widehat{p}_1 - \widehat{p}_2) - \Delta_o}{\sqrt{\widehat{p}\widehat{q}(1/n_1 + 1/n_2)}} \tag{10}$$

donde: $\widehat{p} = \dfrac{x_1 + x_2}{n_1 + n_2}$ $\quad \widehat{q} = 1 - \widehat{p}, \quad \widehat{p}_1 = \dfrac{x_1}{n_1}$ y $\widehat{p}_2 = \dfrac{x_2}{n_2}$

$$F = \frac{s_1^2}{s_2^2} \sim F_{v_1:n_1-1;v_2:n_2-1} \tag{11}$$

Regla1: Si el $EdeP$ cae en la $RdeR$, entonce se rechaza la H_0 y se **acepta** H_a como verdadera. Si por el contrario el $EdeP$ NO cae en la $RdeR$, entonces NO se rechaza H_0, no existe suficiente evidencia para rechazarla, **asumimos** que H_0 es verdad.

Regla2: Si $\alpha >$ valor-p entonces rechazamos H_0, se **acepta** H_a. Si por el contrario $\alpha < valor - p$, no rechazamos H_0, **asumimos** que H_0 es verdad.

Regla3: El $valor - p$ se interpreta como el error que puedo cometer al rechazar H_0, siendo esta verdadera (cometer error tipo I). Si este valor es considerado como pequeño, rechazo H_0, se **acepta** H_a. Si por el contrario se considera este valor grande, entonce no rechazo H_0, asumo que H_0 es verdad.

Tipos de pruebas

$H_0 : \theta = \theta_o$ vs $H_0 : \theta \neq \theta_o$ $\qquad RdeR * |$———————$| * RdeR$

$H_0 : \theta \leq \theta_o$ vs $H_0 : \theta > \theta_o$ $\qquad$ ———————————-$| * RdeR **$

$H_0 : \theta \geq \theta_o$ vs $H_0 : \theta < \theta_o$ $\qquad ** RdeR * |$———————————

Códigos en R pruebas paramétricas

```
# Entrada de datos
x1=c(7, 13, 6, 5, 5, 10, 8, 6, 7)
x2=c(3,7,2,3,6,2,1,0,2)
```

Una población

```
z.test(datos,mu=10,stdev=4, conf.level=98)
t.test(datos, mu=10,conf.level=0.98)
t.test(datos, mu=10,conf.level=0.98,alternative="greater")
t.test(datos, mu=10,conf.level=0.98,alternative="less")
prop.test(x,n, p=0.20, conf.level=0.98)
```

Dos poblaciones

```
t.test(x1,x2,  paired=TRUE)
t.test(x1,x2,  paired=FALSE, var.equal=TRUE, conf.level=0-98)
t.test(x1,x2,  paired=FALSE, var.equal=FALSE, conf.level=0.98)
var.test(x,y)
prop.test(c(x1,x2), c(n1,n2))
```

.

Códigos en R pruebas no paramétricas

```
library(BSDA)
sign.test(x,md=6.5,alternative= "less")  "greater", "two.sided"
sign.test(x,y,md=0, "two.sided")

wilcox.test(x,y,paired=TRUE, laternative="two.sided")
wilcox.test(x,y,paired=FALSE, alternative="two.sided")

runs.test(x)

obs=c(10,15,20)
esp=c(0.333,0.333,0.334)
chisq.test(obs,p=esp)
```

Códigos en R pruebas de normalidad

```
x=rnorm(200,100,25)
hist(x)
qqnorm(x)
qqline(x)

shapiro.test(x)

library(nortest)
ad.test(x)
cvm.test(x)
lillie.test(x)
pearson.test(x)
sf.test(x)

library(normtest)
ajb.norm.test(x)
frosini.test(x)
geary.test(x)
hegazy1.test(x)
hegazy2.test(x)
jb.test(x)
skewness.test(x)
kurtosis.test(x)
dpirgelhalter(x)
wb.test(x)
```

Profesor Daniel Enrique González Gómez
Curso: Probabilidad y Estadística
ver.2.0 oct 25 2021