

## 7. Ejercicios de contraste de hipótesis(parametricos)

### Introducción a la estadística

**Problema 1.** Algunas de las siguientes afirmaciones se refieren a la hipótesis nula, algunas a la hipótesis alternativa.<sup>1</sup> Indique la hipótesis nula,  $H_0$  y la hipótesis alternativa,  $H_1$ , en términos del parámetro apropiado ( $\mu$  o  $p$ ).

- a) El número medio de años que los estadounidenses trabajan antes de jubilarse es 34.
- b) Como máximo, el 60% de los estadounidenses vota en las elecciones presidenciales.
- c) El salario inicial promedio para los graduados de la Universidad Estatal de San José es de al menos 100,000 por año.
- d) El 29% de los estudiantes de último año de secundaria se emborrachan cada mes.
- e) Menos del 5% de los adultos viajan en autobús al trabajo en Los Ángeles.
- f) El número medio de automóviles que posee una persona en su vida no es más de diez.
- g) Alrededor de la mitad de los estadounidenses prefieren vivir lejos de las ciudades, dada la opción.
- h) Los europeos tienen un promedio de vacaciones pagadas cada año de seis semanas.
- i) La probabilidad de desarrollar cáncer de seno es inferior al 11% para las mujeres.
- j) El costo de matrícula promedio de las universidades privadas es de más de 20,000 por año.

**Problema 2.** El precio promedio de los autos medianos en una región es de 32,000 euros. Se realiza una prueba para ver si eso es verdadero. Indique los errores de Tipo I y Tipo II en oraciones completas.

**Problema 3.** Se prueba un saco de dormir para soportar temperaturas de  $-15^\circ\text{F}$ . Cree que la bolsa no puede soportar temperaturas tan bajas. Indique los errores de Tipo I y Tipo II en oraciones completas.

**Problema 4.** En palabras, describa la potencia del test para el ejercicio anterior.

**Problema 5.** Un grupo de médicos está decidiendo si realizar o no una operación. Supongamos que la hipótesis nula,  $H_0$ , es: el procedimiento quirúrgico irá bien. Indique los errores de Tipo I y Tipo II en oraciones completas. Responder:

**Problema 6.** Un grupo de médicos está decidiendo si realizar o no una operación. Supongamos que la hipótesis nula,  $H_0$ , es: el procedimiento quirúrgico irá bien. ¿Cuál es el error con mayor consecuencia?

**Problema 7.** La potencia de una prueba es 0.981. ¿Cuál es la probabilidad de un error tipo II?

**Problema 8.** Un grupo de buzos está explorando un viejo barco hundido. Supongamos que la hipótesis nula,  $H_0$ , es: el barco hundido no contiene tesoros enterrados. Indique los errores de Tipo I y Tipo II en oraciones completas.

**Problema 9.** Un microbiólogo está probando una muestra de agua para E-coli. Supongamos que la hipótesis nula,  $H_0$ , es: la muestra no contiene E-coli. La probabilidad de que la muestra no contenga E-coli, pero el microbiólogo cree que sí es 0.012. La probabilidad de que la muestra contenga E-coli, pero el microbiólogo cree que no contiene 0.002. ¿Cuál es el poder de esta prueba?

---

<sup>1</sup>De Anza College, Introductory Statistics de Barbara Illowsky y Susan Dean

**Problema 10.** Un microbiólogo está probando una muestra de agua para E-coli. Supongamos que la hipótesis nula,  $H_0$ , es: la muestra contiene E-coli. ¿Cuál es el error con mayor consecuencia?

**Problema 11.** Los errores en la conducción se enumeran como la causa de aproximadamente el 54% de todos los accidentes automovilísticos fatales, según la Asociación Americana del Automóvil. Se examinan 30 accidentes fatales seleccionados al azar, y se determina que 14 fueron causados por un error del conductor. Usando  $\alpha = 0,05$ , ¿es precisa la proporción AAA?

- a) Indique la hipótesis nula y alternativa.
- b) Indique el valor p y el valor  $\alpha$
- c) ¿Cuál es su decisión?
- d) Escriba su conclusión

**Problema 12.** Una encuesta reciente en el NY Times indicó que el 48.8% de las familias poseen acciones. Un corredor quería determinar si esta encuesta podría ser válida. Estudió una muestra aleatoria de 250 familias y descubrió que 142 poseían algún tipo de stock. Al nivel de significancia de 0.05.

- a) Indique la hipótesis nula y alternativa.
- b) Indique el valor p y el valor  $\alpha$
- c) ¿Cuál es su decisión?
- d) Escriba su conclusión

**Problema 13.** Según el sitio web del Centro para el Control de Enfermedades, en 2011 al menos el 18% de los estudiantes de secundaria fumaron un cigarrillo. Una clase de Introducción a la Estadística en una población navarra, realizó una prueba de hipótesis en el instituto local (un grupo demográfico de tamaño medio, aproximadamente 1,200 estudiantes, ciudad pequeña) para determinar si el porcentaje de la escuela secundaria local era más bajo. Ciento cincuenta estudiantes fueron elegidos al azar y encuestados. De los 150 estudiantes encuestados, 82 han fumado. Use un nivel de significancia de 0.05 y use evidencia estadística apropiada, realice una prueba de hipótesis y establezca las conclusiones.

- a) Indique la hipótesis nula y alternativa.
- b) Indique el valor p y el valor  $\alpha$
- c) ¿Cuál es su decisión?
- d) Escriba su conclusión

**Problema 14.** La Leche League International informa que la edad promedio para destetar a un niño de la lactancia materna es de cuatro a cinco años en todo el mundo. En Estados Unidos, la mayoría de las madres lactantes destetan a sus hijos mucho antes. Supongamos que se realiza una encuesta aleatoria de 21 madres estadounidenses que recientemente destetaron a sus hijos. La edad media de destete fue de nueve meses (3/4 años) con una desviación estándar de 4 meses. Realice una prueba de hipótesis para determinar si la edad media de destete en los EE. UU. es menor de cuatro años.

**Problema 15.** Las enfermeras dadas de alta recibieron un salario medio de \$69 110. Para ese mismo año, se hizo una encuesta de 41 enfermeras en la región X para poder evaluar si en esa región el salario medio anual era más de \$69 110. La media muestral fue de \$71 121 con una desviación típica de \$7 489. Haz un test de hipótesis para evaluar la cuestión.

**Problema 16.** "Macarrones con queso, por favor !!" por Nedda Misherghi y Rachelle Hall.

Como pobre estudiante hambrienta, no tengo mucho dinero para gastar ni siquiera para las necesidades básicas. Así que mi comida principal favorita y principal es macarrones con queso. Tiene un alto sabor y un bajo costo y valor nutricional.

Un día, mientras me sentaba para determinar el significado de la vida, sentí un antojo serio por esta, tan importante, comida de mi vida. Así que fui a comparar una caja de macarrones con queso, ¡pero era TAN caro! \$ 2.02 !!!

¿Puedes creerlo? Me hizo parar y pensar. El mundo está cambiando rápido. Pensé que el costo promedio de una caja (el tamaño normal, no un paquete de valor familiar supergigante) era como máximo de \$ 1, pero ahora no estaba tan segura. Sin embargo, estaba decidida a averiguarlo.

Fui a 53 de las tiendas de comestibles más cercanas y examiné los precios de los macarrones con queso. Aquí están los datos que escribí en mi cuaderno:

Numero de tiendas	Precio de caja de Mac&Cheese
5	2,02
15	0,25
3	1,29
6	0,35
4	2,27
7	1,50
5	1,89
8	0,75

Pude ver que el costo variaba, pero tuve que sentarme para averiguar si tenía razón o no.

**Problema 17.** El costo de un periódico diario varía de una ciudad a otra. Sin embargo, la variación entre los precios se mantiene estable con una desviación estándar de 20 céntimos. Se realizó un estudio para evaluar si el costo promedio de un periódico diario es de 1.00. Doce costos arrojan un costo promedio de 95 céntimos con una desviación estándar de 18 céntimos. ¿Los datos respaldan la hipótesis al 1%?

**Problema 18.** De generación en generación, la edad promedio en que los fumadores comienzan a fumar varía. Sin embargo, la desviación estándar de esa edad se mantiene constante en alrededor de 2,1 años. Se realizó una encuesta a 40 fumadores de esta generación para ver si la edad promedio de inicio es de al menos 19. La media muestral fue de 18.1 con una desviación estándar muestral de 1.3. ¿Los datos respaldan el reclamo al nivel del 5%?

**Problema 19.** Tenemos los siguientes datos sobre la media del tiempo que un grupo de niños y niñas dedican diariamente a jugar con el balón:

	n	$\bar{x}$	sd
Chicas	9	2 horas	$\sqrt{0,75}$
Chicos	16	3,2 horas	1

a) Haz un test para la diferencia de medias asumiendo que las varianzas son iguales.

b) Haz un test para la diferencia de medias asumiendo que las varianzas son desiguales.

diferencia con varianza conocida

**Problema 20.** Se realizó un estudio para determinar la resistencia a la ruptura de dos tipos de acero. Para una muestra aleatoria formada por 20 especímenes de acero laminado en frío la resistencia promedio  $\bar{x} = 29,8$  ksi. Al estudiar una segunda muestra aleatoria de 25 especímenes de acero galvanizado de dos lados se obtuvo una resistencia promedio muestral  $\bar{y} = 32,7$  ksi. Se supone que las distribuciones de la resistencia a la ruptura de los dos tipos de acero son normales con  $\sigma_x = 4$  y  $\sigma_y = 5$ . Indican los datos que las medias de resistencia a la ruptura son diferentes para los dos tipos de acero?

**Problema 21.** Se tienen las mediciones del nivel de hierro en la sangre de dos muestras de niños: un grupo de niños sanos y el otro padece fibrosis quística. Del primer grupo se tienen 9 mediciones, que dan

$x = 18.9\mu\text{mol/l}$  y  $s_1 = 5,9$ , para el segundo grupo se tienen 13 mediciones que dan  $y = 11,9\mu\text{mol/l}$  y  $s_2 = 6,3\mu\text{mol/l}$ . Estas mediciones proceden de v.a's con distribución normal donde la varianza es la misma. A un 5%, ¿podemos afirmar que las medias son iguales en ambos grupos?

**Problema: para la media con varianza conocida R1.** En "empresa.txt" se proporciona una muestra de 20 salarios procedente de los salarios de los empleados de una multinacional. Se sabe que en esa multinacional, el salario se distribuye normalmente con desviación típica 400. ¿Podría admitirse un valor de 2100 para el salario medio mensual a un nivel de sig. del 0,05%?

**Problema: para la media con varianza conocida R2.** El registro histórico de los últimos tres años muestra que las tasas de accidentes de camiones en las carreteras de 2 carriles en una (ciudad hipotética) de Zreeha es 0.5 accidentes por millón de kilómetros por vehículo con una desviación estándar de 0.1. Después de completar el curso obligatorio de actualización para conductores de camiones, se registraron las tasas de accidentes en 50 sitios aleatorios de carreteras de 2 carriles para estimar las estadísticas actuales. Teniendo en cuenta los datos de la tasa de accidentes, calcule si el curso de manejo hizo alguna diferencia en términos de las tasas de accidentes.

DATOS(Tasa de accidentes): 0.3, 0.91, 0.69, 0.57, 0.28, 0.86, 0.68, 0.36, 0.83, 0.88, 0.76, 0.85, 1.05, 0.7, 0.83, 0.76, 0.51, 0.6, 0.8, 0.71, 1.24, 1.09, 0.23, 0.67, 0.97, 0.97, 0.37, 0.31, 0.76, 0.63, 0.91, 0.97, 0.4, 0.59, 1.05, 0.57, 0.12, 0.69, 0.86, 0.7, 0.6, 0.69, 0.51, 0.72, 1.27, 0.82, 0.43, 0.86, 0.32, 0.83

**Problema: para la media con varianza desconocida R3.** Dados los datos del beneficio neto semestral de diez empresas en "empresa.txt": ¿Existe evidencia en los datos de que el beneficio neto semestral medio es mayor de 41900 euros? Emplee los valores críticos para  $\alpha=7\%$ . ¿Cuál es el p-valor asociado al contraste anterior? Interpretelo.<sup>2</sup>

**Problema: para la media con varianza desconocida R4.** La base de datos "mice" del paquete "datarium", contiene información sobre el peso de diez ratones. ¿Difiere el peso medio de los ratones de 25g?(hacer un t-test para evaluar la cuestión)

**Problema: para la media con varianza desconocida R5.** [Ross] Los datos históricos indican que el nivel medio de acidez (pH) de la lluvia en una determinada región industrial es de 5.2. Para ver si ha habido algún cambio reciente en este valor, los niveles de acidez de 12 tormentas en el pasado año se han medido, con los siguientes resultados:

6.1, 5.4, 4.8, 5.8, 6.6, 5.3, 6.1, 4.4, 3.9, 6.8, 6.5, 6.3

¿Son estos datos lo suficientemente sólidos, con un nivel de significación del 5 %, para que podamos concluir que la acidez de la lluvia ha cambiado desde su valor histórico?

**Problema: para la diferencia de medias varianzas desconocidas y desiguales\* R6.** Se tomaron dos grupos independientes de unidades de un producto fabricadas por la misma máquina. A las unidades del primer grupo se les aplicó un sistema de acabado convencional, y a las del segundo, un sistema nuevo, que parece resultar más caro que el anterior. Los datos correspondientes a los costes en ambos casos se encuentran en "producto.txt". Si los costes de ambos métodos de acabado siguen una distribución normal, ¿existe evidencia de que el coste medio con el acabado nuevo es superior al coste medio con el acabado convencional? ¿Se mantienen los resultados si se sabe que las varianzas de ambos costes son iguales?

**Problema: Para la diferencia de medias varianzas desconocidas y desiguales R7.** Cargando los datos de la librería MASS : `data("birthwt")` sobre el peso de los bebés de madres que fuman y madres que no fuman, ¿podemos decir que hay alguna diferencia entre el peso al nacer de los bebés de madres fumadoras vs. las madres no fumadoras?

**Problema: para la proporción R8.** En futuro.txt se recogen los datos correspondientes a una muestra aleatoria de individuos a los que se les preguntó si tenían una visión optimista del futuro, correspondiendo el valor 1 a aquellos individuos que sí la tenían y el valor 0 al resto. En base a esta información, ¿existe evidencia para rechazar un valor de al menos 0.2 para la proporción de individuos con visión optimista del futuro en la población?

---

<sup>2</sup>Unidad de Innovación, UM