

Complementaria IIND-2107

Ejercicios semana 2

Iván Andrés Trujillo Abella

MIIND

i.trujilloa@uniandes.edu.co

R como calculadora

...

En esta sesión nos apoyaremos de R como calculadora para realizar nuestros cálculos.

...

Vamos a utilizar google colab (**click aquí**) para no tener que descargar nada

Antes de empezar una idea fundamental

Click aquí

Problema 3

Enunciado

Se llevó a cabo un estudio para saber si el aumento en la concentración de sustrato tiene un efecto apreciable sobre la velocidad de una reacción química. Con una concentración de sustrato de 1.5 moles por litro, la reacción se realizó 15 veces, con una velocidad promedio de 7.5 micro moles por 30 minutos y una desviación estándar de 1.5. Con una concentración de sustrato de 2.0 moles por litro, se realizaron 12 reacciones que produjeron una velocidad promedio de 8.8 micro moles por 30 minutos y una desviación estándar muestral de 1.2.

Pregunta

¿Este incremento en la concentración de sustrato ocasiona un aumento en la velocidad media de la reacción de más de 0.5 micromoles por 30 minutos? Utilice un nivel de significancia $\alpha = 0.01$.

Reconozcamos el problema

Cual es nuestro problema en términos generales?

Reconozcamos el problema

Cual es nuestro problema en términos generales?

Si aumento el sustrato incrementa la velocidad de la reacción?

que necesitamos?

Deberíamos comparar la media de la velocidad con menos sustrato y la media de la velocidad con más sustrato.

inferencia

- La diferencia observada es por variabilidad del muestreo?
- Es esta variabilidad mayor a 0.5
- filosofía: Si asumimos que es verdad que la diferencia entre los parámetros es de 0.5 cual es la probabilidad de observar una diferencia mayor de la que nos dan el problema.

t-test

Cuando usamos el t-test debemos reconocer

- las medias están relacionadas o son independientes?
- tamaño de las muestras
- varianzas iguales o no?

Algoritmo

- obtenga el estadístico de prueba
- calcule la probabilidad de observar ese estadístico
- compare la probabilidad anterior con el valor de significancia

Extraíganos la información

Etiquetas a los grupos

Antes-menor concentración = 1, Después-mayor concentración = 2.

\bar{y} variable de interés

- \bar{y}_1 para el grupo 1 y \bar{y}_2 para el grupo 2.

S desviaciones

- S_1 desviación en muestra grupo 1 y S_2 desviación muestra grupo 2.

n tamaño de los grupos

- n_1 para el grupo 1 y n_2 para el grupo 2.

Información

Grupo	1	2
\bar{y}		
S		
n		

Información

Grupo	1	2
\bar{y}		
S		
n		

Grupo	1	2
\bar{y}	7.5	8.8
S	1.5	1.2
n	15	12

Planteemos las hipótesis

Hipótesis

$$H_0 : \mu_2 - \mu_1 = 0.5$$

$$H_1 : \mu_2 - \mu_1 > 0.5$$

(1)

t-test Equations and Degrees of Freedom

Test Type	t-Statistic Equation	Degrees of Freedom
Paired t-test	$t = \frac{\bar{d}}{s_d / \sqrt{n}}$	$df = n - 1$
Unpaired t-test (Equal Variance)	$t = \frac{\bar{X}_1 - \bar{X}_2}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$	$df = n_1 + n_2 - 2$
Unpaired t-test (Unequal Variance, Welch's)	$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$	$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$

Table: uses

in the unpaired test exist the conservative approach of calculate $df = \min(n_1 - 1, n_2 - 1)$.

Tenemos medias no pareadas, asumiendo igual varianzas, tenemos que evaluar la diferencia con respecto a la hipótesis nula de que la diferencia poblacional es de 0.5

Estadístico de prueba

$$E_p = \frac{(\bar{y}_2 - \bar{y}_1) - 0.5}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad (2)$$

Pooled variance

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} \quad (3)$$

...

$$df = n_1 + n_2 - 2 \quad (4)$$

Recuerde la distribución de t

Se distribuye con df grados de libertad.

$$t^*(df, \alpha) \quad (5)$$

aquel valor que bajo la distribución con df grados de libertad la probabilidad de observar un valor extremo es de α porciento.

R calculadora

la CDF nos da

$$CDF(x) = P(X \leq x) \quad (6)$$

```
1 pt(q,df, lower.tail = TRUE) #CDF
2 #q valor al cual queremos hallar la probabilidad.
3 #df los grados de libertad.
4 #lower.tail si queremos computar cola inferior o
  superior.
```

la CDF^{-1} nos da la función cuantíl le pasamos la probabilidad que queremos y nos retorna

$$CDF(CDF^{-1}(P)) = P \quad (7)$$

```
1 qt(p, df, lower.tail = TRUE) # inversa CDF
2 #p es la probabilidad
```

```
1 pt(qt(p = 0.8671, df = df ), df= df)
```

...

Que necesitamos? estudiar cual es la probabilidad de obtener un valor extremo a la derecha del estadístico de prueba.

Computar

Lab click here

De acuerdo a los resultados no hay evidencia suficiente para rechazar H_0 es decir que posiblemente no exista una diferencia poblacional de 0.5.

Problema 4

Enunciado

Se realizó un experimento para determinar la influencia del fármaco sobre la concentración de andrógenos en la sangre. Se obtuvieron muestras de sangre de venados inmediatamente después de recibir una inyección intramuscular. Treinta minutos después se obtuvo una segunda muestra de sangre y después los venados fueron liberados. A continuación, se muestran los datos de la concentración de andrógenos, medida en nanogramos por mililitro (ng/mL), al momento de la captura y 30 minutos más tarde

Pregunta

Pruebe a un nivel de significancia de 10%, si las concentraciones de andrógenos se alteraron después de 30 minutos.

y_pre	y_post
2,76	7,02
1,18	4,12
2,68	5,44
8,1	5,21
7,05	9,26
6,6	8,91
8,79	9,53
7,3	6,85

Table: Dataset

la base de datos la puede conseguir en el siguiente (**click aquí**).

Algoritmo

- Están relacionadas las medias?
- Varianza? (si están o no relacionadas las medias)
- Extraiga la información
- Proponga las hipótesis
- Compute el estadístico de prueba
- Obtenga el valor de P y el valor crítico $t_{(df,\alpha)}$
- Tome una decisión con el valor de P y el nivel de significancia

Laboratorio

Hipótesis

$$\begin{aligned}H_0 : \mu_{diff} &= 0 \\H_1 : \mu_{diff} &\neq 0\end{aligned}\tag{8}$$

Ep

$$t = \frac{\bar{d}}{\frac{s_d}{\sqrt{n}}}\tag{9}$$

$$df = n - 1\tag{10}$$

Laboratorio

(Click aquí)

Decisión

Dado que el valor (absoluto) de el estadístico de prueba no supera el valor crítico $t_{df,\alpha}$ no se rechaza la hipótesis nula, o como el valor de P no es menor a 0.05 entonces no se rechaza la hipótesis nula.