

ESTADÍSTICA APLICADA
Primer Curso del Grado en Bioquímica (2018-2019)

Tema 4: INTERVALOS DE CONFIANZA

4.1. Consideremos los datos `168Perros.txt` del Tema 3. Supongamos que el logaritmo de la alanina aminotransferasa, $\log(\text{ALT})$, sigue una distribución normal. Calcular un intervalo de confianza al 95 % para la media y la varianza de esta distribución.

4.2. Los loci de un marcador genético pueden ser más o menos polimórficos dependiendo del número de alelos y de sus frecuencias poblacionales. El nivel de polimorfismo de un marcador se puede medir mediante la proporción de individuos de la población que son heterocigóticos para ese marcador. En otras palabras, la probabilidad de que un individuo elegido al azar sea heterocigótico se utiliza como medida del grado de polimorfismo.

Supongamos que un investigador encuentra un nuevo polimorfismo de ADN. Se toma una muestra de 50 individuos para medir la heterocigosidad para el marcador correspondiente en la población. Las frecuencias genotípicas aparecen en la siguiente tabla:

Genotipo	1/1	1/2	1/3	2/2	2/3	3/3
Frecuencia absoluta	2	23	2	13	9	1

Determina un estimador puntual y un intervalo de confianza al 95 % para la proporción real p de individuos heterocigóticos. ¿Qué tamaño muestral n habría que haber tomado si hubiéramos querido estimar p con un error menor que 0.1 al nivel de confianza del 95 %?

4.3. Se han tomado las tensiones sanguíneas en una muestra de 10 pacientes hipotensos, obteniéndose las siguientes mediciones:

10 10.5 11 10.7 10.8 12 11.5 9.1 11.3 9.9.

Suponiendo una distribución normal de las tensiones en la población de hipotensos observada, hallar los intervalos de confianza al nivel de confianza del 90 %

- a) para la media μ ;
- b) para la varianza σ^2 de esta población.

4.4. Un microbiólogo quiere comparar dos fármacos, A y B, en términos de sus propiedades antimicrobianas sobre la *Escherichia coli*. La *E. coli* fue subcultivada y luego inoculada en agar Müller-Hinton. Se impregnó un disco con el fármaco A o B para posteriormente colocarlo sobre el agar en medio de la placa, que se incubó a $35 \pm 2^\circ$ durante 24 horas. La máxima anchura de la zona en la que no creció la *E. coli* se midió en 10 placas expuestas a cada uno de los fármacos. A continuación reproducimos las zonas de inhibición (en mm) observadas:

Fármaco A (X)	Fármaco B (Y)
10.5	11.2
7.8	10.3
9.3	9.2
8.7	9.7
10.2	8.9
8.9	10.7
7.4	9.9
9.3	10.1
8.7	7.8
7.9	8.9

Suponemos que $X \sim N(\mu_1, \sigma_1)$ e $Y \sim N(\mu_2, \sigma_2)$.

- a) Calcular estimaciones de las medias μ_i y de las varianzas σ_i^2 .
- b) Calcular un intervalo de confianza para el cociente de varianzas σ_1^2/σ_2^2 al nivel de confianza 0.90. ¿Es razonable suponer que $\sigma_1 = \sigma_2$?
- c) A nivel de confianza $1 - \alpha = 0.95$, calcular un intervalo de confianza para la diferencia de las medias, $\mu_1 - \mu_2$, suponiendo que $\sigma_1 = \sigma_2$. En base al intervalo de confianza obtenido, determinar si hay diferencias entre los fármacos A y B.

4.5. Los siguientes datos proporcionan la recuperación de bromuro adicionado a muestras con contenido vegetal, medido mediante un método cromatográfico gas-líquido. La cantidad de bromuro potásico añadido a cada hortaliza fue la misma.

Tomate	777	790	759	790	770	758	764	$\mu\text{g g}^{-1}$
Pepino	782	773	778	765	789	797	782	$\mu\text{g g}^{-1}$

- a) Hallar un intervalo de confianza para la diferencia de medias. ¿Que hipótesis se están utilizando implícitamente para poder obtener el intervalo?
- b) Determinar si la hipótesis de igualdad de varianzas es razonable.

4.6. La beta-endorfina humana (BEH) es una hormona segregada por la glándula pituitaria bajo condiciones de estrés. Un investigador realizó un estudio para investigar si un programa de ejercicio regular podría afectar a las concentraciones en reposo (sin estrés) de BEH en la sangre. Midió los niveles de BEH en sangre, en enero y de nuevo en mayo, de 10 participantes en un programa de ejercicio físico. Los resultados fueron

Participante	Enero	Mayo
1	42	22
2	47	29
3	37	9
4	9	9
5	33	26
6	70	36
7	54	38
8	27	32
9	41	33
10	18	14

Construir un intervalo de confianza al 95% para la diferencia de niveles esperados de BEH entre enero y mayo. ¿Existe evidencia de que los niveles de BEH son menores en mayo que en enero?

4.7. Según un ensayo fiable, el contenido de ATP (trifosfato de adenosina) de un cierto tipo de célula es $111 \mu\text{mol}/100 \text{ mL}$. Al diseñar un nuevo ensayo se obtuvieron los siguientes valores en análisis replicados: 117, 119, 111, 115, $120 \mu\text{mol}/100 \text{ mL}$ (promedio = 116.4). ¿Se puede tener una confianza del 95% de que ese resultado difiere del valor “conocido”? Especificar las hipótesis que se emplean en la solución del problema.

4.8. En el paquete `PairedData` de R aparece el conjunto de datos `Shoulder`. Se trata de la flexibilidad (medida en grados según el rango de movilidad del brazo) de los hombros izquierdo y derecho en un grupo de 15 nadadores (`Swimmer`) y en otro grupo de 15 personas sedentarias (`Control`). Para reproducir los datos con R, ejecutar el código

```
library("PairedData")
Shoulder
```

Algunos cálculos realizados sobre estas observaciones:

```
Nadador = (Shoulder$Group=="Swimmer")
```

```
HombrosNadador=Shoulder[Nadador,c("Left","Right")]
HombrosControl=Shoulder[!Nadador,c("Left","Right")]
```

```
mean(HombrosNadador$Left)
[1] 201.8
mean(HombrosNadador$Right)
[1] 200.9333
mean(HombrosControl$Left)
[1] 185.6
mean(HombrosControl$Right)
[1] 185.4667
var(HombrosNadador$Left)
[1] 28.45714
var(HombrosNadador$Right)
[1] 39.49524
var(HombrosControl$Left)
[1] 108.1143
var(HombrosControl$Right)
[1] 97.69524
```

- Calcular un intervalo de confianza al nivel 0.95 para la diferencia entre las flexibilidades esperadas del hombro derecho e izquierdo en un nadador. Especifica las suposiciones previas para la resolución.
- Determina un intervalo de confianza al 95 % para la flexibilidad esperada en el hombro derecho de un nadador. Especifica las suposiciones previas para resolver este apartado.
- Calcular un intervalo de confianza al 95 % para la diferencia entre las flexibilidades esperadas del hombro derecho en nadadores y controles.
- Explica qué hace el siguiente código e interpreta el correspondiente output de R:

```
boxplot(Shoulder$Left~Shoulder$Group,main="Left")
boxplot(Shoulder$Right~Shoulder$Group,main="Right")
```

