

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

Tema 5: CONTRASTES DE HIPÓTESIS

5.1. La enantiofilia es un fenómeno por el cual el pistilo de una flor está desviado a la izquierda o a la derecha, de manera que cada tipo de flor es la imagen especular de la otra. Esta asimetría promueve la polinización cruzada en plantas visitadas por abejas, y disminuye el nivel de autopolinización. La orientación (a izquierda o derecha) de una planta describe la localización de su pistilo (ver Figura 1). La orientación a izquierda es un rasgo recesivo. Bajo cierto modelo genético, una flor de segunda generación resultante del cruce de dos homozigotos (uno dominante y otro recesivo) tiene una probabilidad $p = 0.25$ de estar orientada a la izquierda. En un experimento de este tipo, en la segunda generación se obtuvieron un total de 9 flores orientadas a la izquierda y 24 orientadas a la derecha.

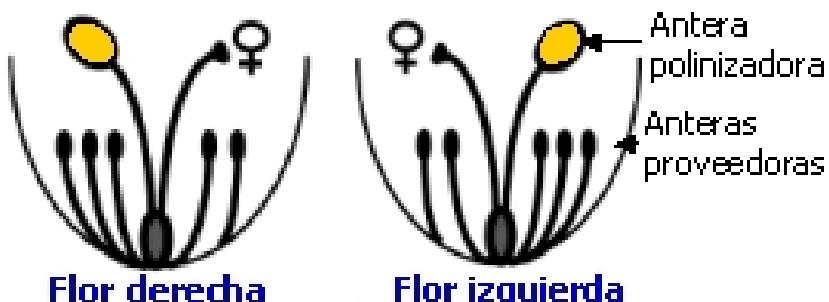


Figura 1

- En base a la muestra observada obtener un intervalo de confianza al 95 % para p .
- A un nivel de significación de 0.05, ¿hay suficiente evidencia en contra de la hipótesis de que $p = 0.25$?
- Explica el siguiente código de R e interpreta la correspondiente salida:

```
prop.test(9,33,p=0.25,alternative="less",correct=FALSE)

1-sample proportions test without continuity correction

data: 9 out of 33, null probability 0.25
X-squared = 0.090909, df = 1, p-value = 0.6185
alternative hypothesis: true p is less than 0.25
95 percent confidence interval:
 0.0000000 0.4137471
sample estimates:
      p 
0.2727273
```

5.2. Los contaminantes provenientes de aguas residuales y vertidos industriales pueden reducir la concentración de oxígeno disuelto en el agua y tener efectos adversos en las especies acuáticas. Se obtienen medidas de la concentración (en ppm) de O₂ en diversas localizaciones de un río:

4.9 5.1 5.6 4.3 4.7 4.9 4.5 5.1

Supongamos que la concentración de O₂ en el agua sigue una distribución normal de media μ y desviación típica σ .

- a) Calcular un intervalo de confianza al 95 % para μ .
- b) Calcular un intervalo de confianza al 90 % para σ^2 .
- c) Algunos científicos piensan que 5.0 ppm es una concentración de O₂ disuelto en la que apenas pueden vivir los peces. En base a la muestra observada, al nivel de significación 0.05, determinar si hay suficiente evidencia de que la concentración media de oxígeno disuelto es menor que 5.0 ppm.

5.3. El siguiente conjunto de datos procede del National Quality Control Scheme, Queen Elizabeth Hospital, Birmingham¹. Son los resultados de la cantidad de sodio (en ppm) presente en 20 muestras procedentes de análisis séricos:

140	143	141	137	132	157	143	149	118	145
138	144	144	139	133	159	141	124	145	139

- a) Construye un intervalo de confianza para la cantidad esperada de sodio en una muestra de sangre.
- b) ¿Hay evidencia de que el nivel medio de sodio en sangre es diferente de 140 ppm?

5.4. En un estudio² sobre absorción y metabolismo de vitaminas, se administró a 20 ratas vitamina A durante un período de tres días. A diez de dichas ratas se les proporcionó la vitamina A adicionada a aceite de maíz y a las otras diez se las alimentó con la vitamina incorporada a aceite de ricino. Las ratas fueron sacrificadas el cuarto día y se midió la concentración de vitamina A en el hígado (ver los datos resultantes en la Tabla 1).

Maíz	407	433	458	461	477	491	525	470	486	529
Ricino	382	387	389	433	435	453	461	428	474	486

Tabla 1

Un análisis de los datos con R proporciona los siguientes resultados:

```
Datos = read.table("oilvit.txt", header=TRUE)
Aceite = Datos$type
Maiz = (Aceite == "corn")
X = Datos[Maiz,2]
Y = Datos[!Maiz,2]
mean(X)
[1] 473.7
mean(Y)
[1] 432.8
var(X)
[1] 1408.678
var(Y)
[1] 1372.844
```

¹Fuente de los datos: Andrews, D.F. and Herzberg, A.M. (1985). *Data. A Collection of Problems from Many Fields for the Student and Research Worker*. Springer.

²Bliss, C.I. (1967). *Statistics in Biology*. McGraw-Hill.

- a)** A un nivel de significación del 10 %, ¿hay suficiente evidencia muestral en contra de la igualdad de varianzas entre el grupo alimentado con aceite de maíz y aquél alimentado con aceite de ricino? Especificar las suposiciones previas que garantizan la validez del procedimiento empleado.
- b)** A un nivel del 5 %, ¿proporcionan los datos suficiente evidencia a favor de que la concentración media de vitamina A en el hígado es mayor cuando la alimentación es a base de aceite de maíz que a base de aceite de ricino? ¿Es el p-valor del contraste mayor o menor que 0.05? ¿Y mayor o menor que 0.01? Especificar las suposiciones previas que garantizan la validez del procedimiento empleado.
- c)** Explica qué hace el siguiente código R y determina el valor de las cantidades sustituidas por interrogantes:

```
t.test(X,Y,var.equal=TRUE)
```

```
Two Sample t-test

data: X and Y
t = ??????, df = ??, p-value = 0.02462
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
????????? ????????
sample estimates:
mean of x mean of y
473.7     432.8
```

5.5. La variación en el éxito reproductivo es determinante para el tamaño de la población. En un artículo de *Genetics Research*³ se observó, para un conjunto de 38 *Drosophila* macho marcados con pintura, el número X de apareamientos de cada uno de estos machos durante un período de 14 horas desde que se los introducía en un vial en el que había un cierto número de moscas no marcadas de ambos sexos. Los datos obtenidos aparecen resumidos en la Tabla 2. Una de las conclusiones del artículo era que, bajo ciertas condiciones experimentales, la variable X seguía una distribución de Poisson de parámetro $\lambda > 0$.

Número de apareamientos por macho	Frecuencia observada
0	12
1	8
2	11
3	3
4	4

Tabla 2

- a)** Calcula $\hat{\lambda}$, la estimación de máxima verosimilitud de λ para esta muestra.
- b)** A nivel de significación 0.05, ¿hay evidencia muestral en contra del modelo Poisson?

³Joshi, A., Do, M.H., Mueller, L.D. (1999). Poisson distribution of male mating success in laboratory populations of *Drosophila melanogaster*. *Genet. Res., Camb.*, 73, 239–249.

5.6. Se realiza un cruzamiento prueba de un posible dihíbrido (un doble heterocigoto) de *Drosophila B/b;F/f* con doble homocigoto recesivo *b/b;f/f* (*B* = cuerpo negro, *b* = cuerpo marrón, *F* = quetas dobles, *f* = quetas sencillas). Los resultados son los siguientes:

Cuerpo negro, quetas dobles	230
Cuerpo negro, quetas sencillas	210
Cuerpo marrón, quetas dobles	240
Cuerpo marrón, quetas sencillas	250

Utilizar un contraste χ^2 para determinar si estos resultados se ajustan a los esperados para un cruzamiento prueba de este hipotético dihíbrido.

Observación: Este tipo de experimento que cruza un dihíbrido con un doble homocigoto recesivo y luego aplica un test χ^2 de bondad de ajuste se utiliza para comprobar si dos genes se segregan independientemente.

5.7. Un veterinario midió la densidad de células nerviosas en lugares específicos del intestino de nueve caballos. Los resultados del lugar I (región media del yeyuno) y del lugar II (región mesentérica del yeyuno) se muestran en la tabla siguiente:

Animal	Lugar I	Lugar II
1	50.6	38.0
2	39.2	18.6
3	35.2	23.2
4	17.0	19.0
5	11.2	6.6
6	14.2	16.4
7	24.2	14.4
8	37.4	37.6
9	35.2	24.4

¿Hay suficiente evidencia muestral de que en toda la población de caballos la densidad de células nerviosas del lugar I es superior a la del lugar II?.

5.8. La tabla siguiente muestra los efectos de un placebo y de la hidroclorotiacida sobre la presión sistólica de 12 pacientes, a la mitad de los cuales se les administró el placebo. A los otros seis se les dio la medicación. Supongamos que la presión sistólica es una v.a. con distribución normal.

Placebo	211	210	210	203	196	190
H-cloro	181	172	196	191	167	161

- a) Dar la estimación de máxima verosimilitud de la presión sanguínea sistólica media con ambos tratamientos.
- b) ¿Puede aceptarse que las varianzas, con placebo y con hidroclorotiacida, son iguales?
- c) Asumiendo igualdad de varianzas, determinar si hay suficiente evidencia estadística, a nivel $\alpha = 0.05$, para afirmar que el tratamiento con hidroclorotiacida reduce la presión sanguínea sistólica.
- d) Decidir razonadamente si el p-valor del contraste anterior es superior o inferior a 0.05.

5.9. Se usan el método de la glucosa oxidasa y el de la hexoquinasa para analizar el nivel de glucosa en diez muestras de sangre, obteniéndose los siguientes resultados (en mM):

Muestra	glucosa oxidasa	hexoquinasa
1	1.1	0.9
2	2.0	2.1
3	3.2	2.9
4	3.7	3.5
5	5.1	4.8
6	8.6	8.7
7	10.4	10.6
8	15.2	14.9
9	18.7	18.7
10	25.3	25.0

Determina si hay diferencias entre el nivel medio de glucosa determinada con uno y otro método. ¿Qué hipótesis se utilizan? ¿Depende la conclusión de que se utilicen los niveles de significación $\alpha = 0.01$, $\alpha = 0.05$ y $\alpha = 0.1$? ¿Qué se puede decir del p-valor del contraste?

5.10. Se sospecha que ciertos riesgos ambientales y laborales pueden alterar la proporción de nacimientos de varones. En particular, hay controversia respecto al efecto que la exposición a radiación ionizante tiene sobre el *sex ratio* o cociente del número de nacimientos masculinos respecto a aquél de nacimientos femeninos. Scherb *et al.* (2013)⁴ estudiaron el efecto sobre el *sex ratio* del accidente en la Central Nuclear de Chernóbil (Ucrania) en abril de 1986. Concretamente, desde 1959 hasta 1986 (ambos años incluidos) nacieron en Rusia 31 706 752 varones de un total de 61 890 823 nacimientos. Desde 1987 hasta el 2010 (ambos años incluidos) en Rusia nacieron 38 360 531 bebés de los que 19 727 605 fueron varones.

a) A nivel $\alpha = 0.05$, ¿hay suficiente evidencia estadística de que el accidente de Chernóbil redujo la proporción de nacimientos de niñas (sobre el total de nacimientos) en Rusia?.

b) Explica el siguiente código de R e interpreta la correspondiente salida:

```
prop.test(c(31706752,19727605),c(61890823,38360531),alternative="less",
          correct=FALSE)
```

```
2-sample test for equality of proportions without continuity correction

data: c(31706752, 19727605) out of c(61890823, 38360531)
X-squared = 366.74, df = 1, p-value < 2.2e-16
alternative hypothesis: less
95 percent confidence interval:
 -1.000000000 -0.001798013
sample estimates:
 prop 1    prop 2
0.5123013 0.5142683
```

⁴Fuente: Scherb, H. Kusmierz, R. y Voigt, K. (2013). Increased sex ratio in Russia and Cuba after Chernobyl: a radiological hypothesis. *Environmental Health*, 12:63.