Inferencia Estadística

Test multinomial, bondad de ajuste, test no paramétricos y comparaciones múltiples

San Andrés

June 14, 2022

Test multinomial

Sea $(X_1, \dots X_k) \sim \mathcal{M}(n, \vec{p})$ es decir un vector multinomial. El estimador de máxima verosimilitud es

$$\widehat{p}_{mv} = (\widehat{p}_1, \dots \widehat{p}_k) = (\frac{X_1}{n}, \dots \frac{X_k}{n})$$

Queremos testear

$$H_0 : p = p_0 \quad vs \quad H_1 : p \neq p_0.$$

Test multinomial

Estadístico χ^2 de Pearson

$$T_n = \sum_{j=1}^{k} \frac{(X_j - E_j)^2}{E_j}$$

Donde $E_j = E(X_j) = np_{0j}$.

Bajo H_0 si n es grande $T_n \approx \chi^2_{k-1}$ por lo tanto

$$\text{p-valor} \approx P(\chi^2_{k-1} > T_{obs})$$

Test multinomial

Decidir si un dado está cargado

Si lanzamos n veces un dado y
$X_i =$ cantidad de veces que salió el numero i entonces el vector

$$(X_1, \dots X_6) \sim \mathcal{M}(n, p).$$

- Armar el estadístico de Pearson para decidir si un dado esta cargado.
- Sortear $n = 25$ lanzamientos de dados justos y evaluarlo
- Calcular el p-valor de la muestra.
- Mediante una simulación determinar la potencia del test en los dados cargados definidos en la actividad 1.

Test multinomial

El test multinomial tiene muchas aplicaciones

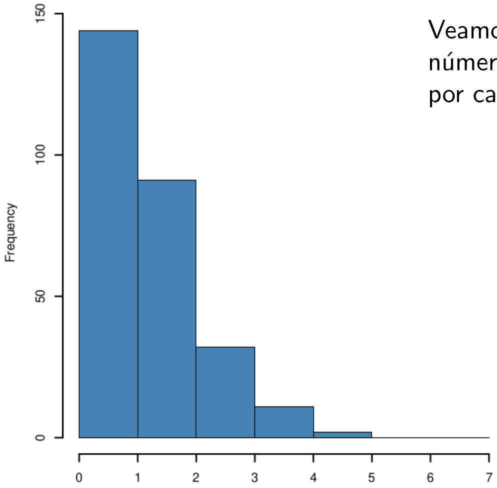
Bondad de ajuste

¿Cómo puedo testear si mis datos vienen de un determinado modelo paramétrico? e.g

Dada una muestra $X_1, \ldots, X_n$

$$H_0 : X \sim Poi(\lambda) \quad vs \quad H_1 : X \not\sim Poi(\lambda).$$

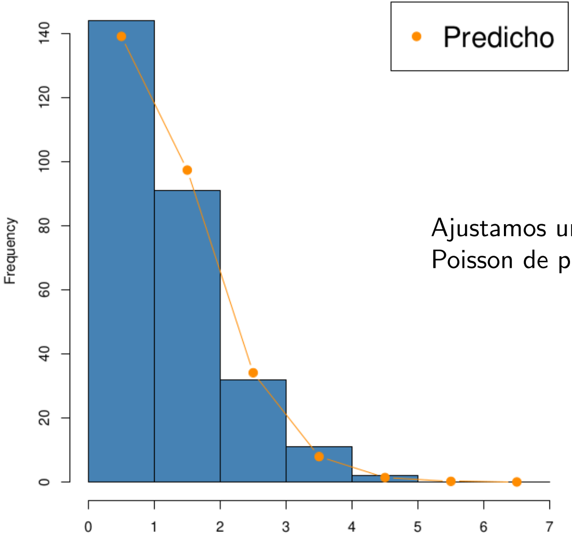
Ejemplo: Patadas de caballo



Veamos la distribución del número de soldados fallecidos por caballería y por año.

0	144
1	91
2	32
3	11
4	2
5	0
6	0
7	0

Ejemplo: Patadas de caballo



Ajustamos una distribución Poisson de parámetro $\lambda = 0.7$.

Bondad de ajuste

Dependiendo del modelo pueden existir test mas potentes que el multinomial.

Kolmogorov Smirnov

Para una muestra $X_1, \ldots X_n$ El test de Kolmogor Smirnov permite testear

$$H_0 : X \sim F \quad vs \quad H_1 : X \nsim F$$

El estadístico del test es

$$D_n = \sup_x |F_n(x) - F(x)|$$

cuya distribución es conocida bajo H_0.

Analisis de la varianza

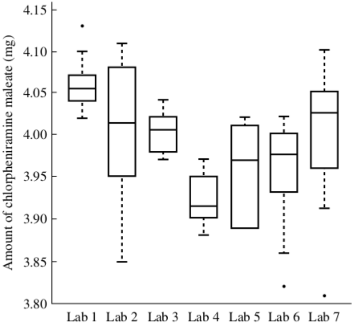
Análisis de la varianza (ANOVA) es una metodología para comparar la media de distintos tratamiento. Esta técnica es una generalización de las técnicas de comparación de dos medias.
La siguiente data es de Kirchhoefer (1979) quien estudió las mediciones de *chlorpheniramine maleate* en tabletas. Se cuenta con las mediciones de comprimidos con valor nomial de 4 mg hechas en 7 laboratorios. Cada uno realizó 10 mediciones.

Hay dos fuentes de posible variación en los datos:

- Variación dentro de cada laboratorio. (Within)
- Variación entre laboratorios. (Between)

Análisis de la varianza

Lab 1	Lab 2	Lab 3	Lab 4	Lab 5	Lab 6	Lab 7
4.13	3.86	4.00	3.88	4.02	4.02	4.00
4.07	3.85	4.02	3.88	3.95	3.86	4.02
4.04	4.08	4.01	3.91	4.02	3.96	4.03
4.07	4.11	4.01	3.95	3.89	3.97	4.04
4.05	4.08	4.04	3.92	3.91	4.00	4.10
4.04	4.01	3.99	3.97	4.01	3.82	3.81
4.02	4.02	4.03	3.92	3.89	3.98	3.91
4.06	4.04	3.97	3.90	3.89	3.99	3.96
4.10	3.97	3.98	3.97	3.99	4.02	4.05
4.04	3.95	3.98	3.90	4.00	3.93	4.06



Análisis de la varianza

Teoría normal - Test F

- Tenemos I grupos cada uno con J observaciones. En el ejemplo $I = 7$, $J = 10$. Los grupos se los suele llamar tratamientos o niveles.
- Llamamos Y_{ij} a la observación j del tratamiento i. El modelo estadístico es
$$Y_{ij} = \mu_i + \varepsilon_{ij}.$$
- Los errores se asumen independientes, normales con media cero y varianza σ^2.

Análisis de la varianza

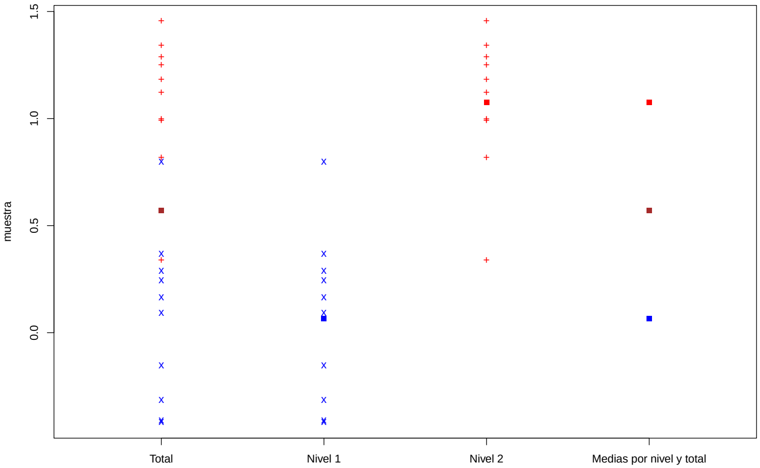
La siguiente identidad es la clave de la metodología:

$$\underbrace{\sum_{i=1}^{I}\sum_{j=1}^{J}\left(Y_{ij}-\overline{Y}_{..}\right)^2}_{S_{tot}} = \underbrace{\sum_{i=1}^{I}\sum_{j=1}^{J}\left(Y_{ij}-\overline{Y}_{i.}\right)^2}_{SS_w} + \underbrace{J\sum_{i=1}^{I}\left(\overline{Y}_{i.}-\overline{Y}_{..}\right)^2}_{SS_b}.$$

Con

$$\overline{Y}_{..} = \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J} Y_{ij} \quad \overline{Y}_{i.} = \frac{1}{J}\sum_{j=1}^{J} Y_{ij}.$$

ANOVA



muestra
1.5
1.0
0.5
0.0
Total
Nivel 1
Nivel 2
Medias por nivel y total

Análisis de la varianza

Teorema

Bajo la hipótesis nula $H_0 : \mu_1 = \mu_2 = \ldots = \mu_I$

- $$\frac{SS_w}{\sigma^2} \sim \chi^2_{I(J-1)}$$
- $$\frac{SS_b}{\sigma^2} \sim \chi^2_{I-1}$$
- SS_w y SS_b son independientes
- $$F = \frac{SS_b}{I-1}\frac{I(J-1)}{SS_w} \sim F_{I-1,I(J-1)}$$

  Donde $F_{I-1,I(J-1)}$ es una distribución F de Fisher con grados de libertad $I-1$ e $I(J-1)$.

Análisis de la varianza

Test de hipótesis

Para realizar un test de nivel α con

$$H_0 : \mu_1 = \mu_2 = \ldots = \mu_I \quad vs \quad \mu_i \neq \mu_k \text{ para algún i, k.}$$

- Región de rechazo

$$\mathcal{R} = \{F > F_{I-1,I(J-1),1-\alpha}\}$$

- P-valor: $P(F_{I-1,I(J-1)} > F_{obs})$.

Test no paramétricos

Tipos de test

Existen distintos tipos de test no paramétricos cada uno con sus ventajas y desventajas

- Test de signos.
- Test de rangos.
- Test de permutaciones.

Test de signos

Test de signos para la mediana de una población

Sea $X_1, \ldots X_n$ una muestra aleatoria de variables continuas. Si llamamos m a la mediana de X, queremos contrastar

$$H_0 : m = m_0 \qquad \text{versus} \qquad H_1 : m \neq m_0.$$

- Llamamos $D_i = \mathbb{I}_{X_i - m_0 > 0}$ bajo H_0 $D_i \sim \mathcal{B}_e(0.5)$.
- Podemos hacer un test binomial ya que

$$X = \sum_{i=1}^{n} D_i \overset{H_0}{\sim} \mathcal{B}(n, 0.5).$$

- $$\text{p-valor} = P\left(|X - \frac{n}{2}| \geq |X_{obs} - \frac{n}{2}|\right).$$

Test de signos

Ejercicio

Sean $(X_1, Y_1), \ldots (X_n, Y_n)$ una muestra aleatoria de variables continuas. Contrastar

$$H_0 : P(X \geq Y) = P(X \leq Y) \quad \text{vs} \quad H_1 : P(X \geq Y) \neq P(X \leq Y).$$

Test de rangos

Test de Man-Whitney

Sean $X_1, \ldots X_n$ e $Y_1, \ldots Y_m$ dos muestras aleatorias independientes con funciones de distribución F_X y F_Y. Supongamos que la distribución de Y es un corrimiento de la de X. Es decir exite $a \in \mathbb{R}$

$$F_Y(x) = F_X(x - a) \quad \forall x \in \mathbb{R}.$$

El test de Man - Whitney permite testear

$$H_0 : a = 0 \qquad \text{vs} \qquad H_1 : \begin{cases} a & \neq 0 \\ a & > 0 \\ a & < 0 \end{cases}.$$

Test de rangos

Observaciones

- Las muestras pueden ser de distinto tamaño ($n \neq m$).
- El caso en que X es normal ya lo estudiamos.
- Si $W = U + a \implies F_W(x) = F_U(x - a)$.

Test de rangos

Construcción del estadístico

1. Agrupar las $n+m$ variables aleatorias como si fueran una única muestra

$$Z_1 = X_1, \ldots Z_n = X_n, Z_{n+1} = Y_1, \ldots Z_{n+m} = Y_m.$$

2. Ordenarlas de menor a mayor

$$Z_{(1)} \leq Z_{(2)} \leq \ldots \leq Z_{(n+m)}.$$

3. Asignamos rango a los elementos de la muestra a partir del orden que les tocó.
4. Sumamos los rangos de la muestra de menor tamaño. A esta cantidad aleatoria la llamamos R.
5. Si este número es 'muy chico" o " muy grande" rechazamos la hipótesis nula.

Test de rangos

Ejemplo

Con $n = 3$ y $m = 2$ los posibles valores de R son $\{3, 4, 5, 6, 7, 8, 9\}$.
Si obtenemos la siguiente muestra

	X_1	X_2	X_3	Y_1	Y_2
Observado	1.5	0.8	5.4	2.1	0.3
Rango	3	2	5	4	1

Obtenemos

$$R_{obs} = 4 + 1 = 5.$$

Test de rangos

Test de Man-Whitney

Pregunta: ¿Qué significa que R sea grande o chico?

- R es una variable aleatoria discreta.
- Bajo H_0 $F_X = F_Y$ por lo que cualquiera de los $(n+m)!$ ordenamientos de los rangos posible es equiprobabile.
- Esto nos dice que la distribución de R bajo H_0 no depende de quien es F_X (es un estadístico libre o pivot).
- La distribución de R está tabulada y el test tiene un comando en el lenguaje R.

Test de rangos

Test de Man-Whitney

Pregunta: ¿Se pueden cambiar algunos supuestos sobre las distribuciones?

- Si se agrega la hipótesis de que F_X y F_Y son simétricas con esperanza finitas, el test de Mann- Whitney se convierte en un test de medias.
- Si no se cumple $F_Y(x) = F_X(x-a)$ el test se puede usar igual pero para las siguientes hipótesis

$$H_0 : F_X = F_Y \quad \text{vs} \quad H_1 : \begin{cases} P(X > Y) \neq P(Y > X) \\ P(X > Y) < P(Y > X) \\ P(X > Y) > P(Y > X) \end{cases}$$

Test no paramétricos

Ejemplo

Analizar los datos de temperatura.

Test de permutaciones

Se tiene $X_1, \dots X_n \sim F_X$ e $Y_1, \dots Y_m \sim F_Y$. Se quiere testea

$$H_0 : F_X = F_Y \quad \text{vs} \quad F_X \neq F_Y.$$

- Se considera un estadistico $T(x_1, \dots, x_n, y_1, \dots, y_m)$ de la muestra, por ejemplo

  $$T(X_1, \dots X_n, Y_1, \dots Y_m) = |\overline{X}_n - \overline{Y}_m|.$$

  O también

  $$T(X_1, \dots X_n, Y_1, \dots Y_m) = |\text{med}\{X_1, \dots X_n\} - \text{med}\{Y_1, \dots Y_m\}|.$$

- Si $N = n + m$ se consideran todas las $N!$ permutaciones del vector $(X_1, \dots X_n, Y_1, \dots Y_m)$ y para cada una computar el estadístico. Obtenemos $T_1, \dots T_{N!}$.

Test de permutaciones

- Bajo la hipótesis nula cada T_j tiene probabilidad $1/N!$ de ocurrir. Obtenemos así la distribución de permutaciones de T.
- Si llamamos t_{obs} el valor observado del estadístco obtenemos

$$p - valor = P_{H_0}(T > t_{obs}) = \frac{1}{N!}\sum_{j=1}^{N!} I(T_j > t_{obs}).$$

- La elección del estadístico T influye en el pvalor.

Test de permutaciones

Ejemplo

Supongamos $n = 2$ y $m = 1$ obtenemos la muestra $(X_1, X_2, Y_1) = (1, 9, 3)$. Si $T(X_1, X_2, Y_1) = |\overline{X} - \overline{Y}|$, entonces $t_{obs} = |5 - 3| = 2$.

Permutación	Valor de T	Probabilidad
$(1, 9, 3)$	2	$1/6$
$(9, 1, 3)$	2	$1/6$
$(1, 3, 9)$	7	$1/6$
$(3, 1, 9)$	7	$1/6$
$(3, 9, 1)$	5	$1/6$
$(9, 3, 1)$	5	$1/6$

El pvalor es $P_{H_0}(T > 2) = 4/6$

Test de permutaciones

En general no es práctico evaluar las $N!$ permutaciones. Lo que se suele hacer es aproximar el p-valor tomando una muestra de posibles permutaciones.

1. Calcular el valor observado del estadístico T

$$t_{obs} = T(X_1, \ldots X_n, Y_1, \ldots Y_m)$$

2. Permutar las variables de manera aleatoria y calcular el estadistico
3. Repetir el paso previo B veces y llamar $T_1, \ldots T_B$ a los valores obtenidos.
4. El pvalor aproximado es

$$\frac{1}{B}\sum_{j=1}^{B} I(T_j > t_{obs}).$$

Test de permutaciones

Example

Calcular el test de permutaciones

Compraciones múltiples

En algunas situaciones se quiere realizar muchos test de hipotesis en simultaneo. Si cada test se conduce a nivel α la probabilidad de rechazar erroneamente H_0 es α para cada uno de los test, pero la probabilidad de que haya al menos una hipótesis nula rechazada incorrectamente es mucho mayor.
Este problema surge en situaciones donde uno testea miles o hasta millones de hipótesis en simultaneo. Hay varias formas de lidiar con este problema. Vamos a discutir dos

- El método de Bonferroni.
- Controlar la tasa de falsos descubrimientos o False Discovery Rate (FDR)

Comparaciones múltiples

Método de Bonferroni

Consideramos m hipótesis a testear

$$H_{0i} \quad \text{vs} \quad H_{1i}, \quad i = 1, \ldots m.$$

Llamemos $P_1, \ldots, P_m$ los p-valores para cada uno de los tests.

Se decide rechazar H_{0i} si

$$P_i < \frac{\alpha}{m}.$$

Teorema

Con el método de Bonferroni la probabilidad de rechazar incorrectamente alguna de las hipótesis nulas es menor o igual a α.

Comparaciones múltiples

El problema de Bonferroni es que si $m = 1000$ y $\alpha = 0.05$ solo rechazamos las hipótesis nulas cuyo p-valor es menor o igual a 0.00005. Lo cual es demasiado conservativo para la mayoría de las aplicaciones prácticas.
En muchas ocasiones tiene mas sentido contralar el **False Discovery Rate** (FDR) que se define como la esperanza de las falsas rechazos divido por el total de rechazos.

Comparaciones múltiples

Supongamos que tenemos un criterio para decidir que hipótesis nulas rechazar. Podemos definir las siguientes variables

	H_0 no rechazada	H_0 rechazada	Total
H_0 verdadera	U	V	m_0
H_0 falsa	T	S	m_1
Total	m -R	R	m

Luego llamamos

$$FDP = \begin{cases} V/R & \text{si } R > 0 \\ 0 & \text{si } R = 0. \end{cases}$$

Y

$$FDR = E(FDP).$$

Comparaciones múltiples

El método Benjamini-Hochberg (BH)

1. Sean $P_{(1)} \leq \ldots \leq P_{(m)}$ los p-valores ordenados.
2. Sean
$$l_i = \frac{i\alpha}{C_m m}, \quad y \quad R = \max\left\{i : P_{(i)} < l_i\right\}.$$
Donde C_m es una constante que vale 1 si los p-valores son independientes.
3. Sea $T = P_{(R)}$ llamamos T al umbral de rechazo de BH.
4. Rechazar todas las hipótesis nulas H_{0i} tales que $P_i \leq T$.

Comparaciones múltiples

Teorema

El procedimiento de BH garantiza que sin importar la cantidad de hipótesis nulas verdaderas ni la distribución de los pvalores cuando la hipótesis nula es falsa

$$FDR = E(FDP) \leq \frac{m_0}{m}\alpha \leq \alpha.$$

Observar que los test no deben ser independientes si se usa la constante $C_m = \sum_{j=1}^{m} \frac{1}{j}$.

Comparaciones múltiples

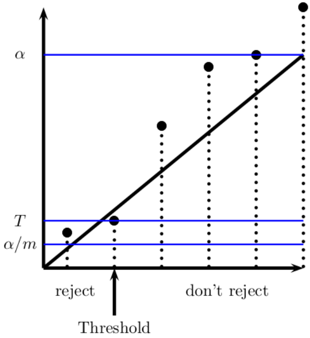


FIGURE 10.6. The Benjamini-Hochberg (BH) procedure. For uncorrected testing we reject when $P_i < \alpha$. For Bonferroni testing we reject when $P_i < \alpha/m$. The BH procedure rejects when $P_i \leq T$. The BH threshold T corresponds to the rightmost undercrossing of the upward sloping line.