Probabilitat i Estadística 1

Problemes Tema 6. Estimació puntual

1. Siguin $X_1, \ldots, X_n$ v.a.i.i.d. amb esperança μ i variància σ^2. Es defineix la variància mostral

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2.$$

Proveu que $\mathbb{E}(S^2) = \sigma^2$.

Resolució:

$$\begin{aligned}
\mathbb{E}(S^2) =& \frac{1}{n-1} \mathbb{E}\left(\sum_{i=1}^{n} (X_i - \bar{X})^2 \right) \\
=& \frac{n}{n-1} \mathbb{E}\left((X_1 - \bar{X})^2 \right) \\
& (\mathbb{E}(X_1 - \bar{X}) = 0) \\
=& \frac{n}{n-1} \mathrm{Var}(X_1 - \bar{X}) \\
=& \frac{n}{n-1} \left(\mathrm{Var}(X_1) + \mathrm{Var}(\bar{X}) - 2\mathrm{Cov}(X_1, \bar{X}) \right) \\
& \left(\mathrm{Cov}(X_1, \bar{X}) = \frac{1}{n} \mathrm{Cov}(X_1, \sum_{j=1}^{n} X_j) = \frac{1}{n} \mathrm{Cov}(X_1, X_1) = \frac{\sigma^2}{n} \right) \\
=& \frac{n}{n-1} \left(\sigma^2 + \frac{\sigma^2}{n} - 2\frac{\sigma^2}{n} \right) \\
=& \frac{n}{n-1} \frac{(n-1)\sigma^2}{n} \\
=& \sigma^2.
\end{aligned}$$

2. Sea $X_1, \ldots, X_n$ una m.a.s. de $X \sim f(x, \lambda) = \lambda e^{-\lambda x + 1}$, $x > 1/\lambda$. Hallar el estimador de momentos de λ.

Resultat: $\hat{\lambda} = 2/\bar{X}_n$

Resolució:
Empezamos calculando el primer momento de X:

$$\mu_1 = \mathbb{E}(X) = \int_{1/\lambda}^{\infty} x \lambda e^{-\lambda x + 1} dx = e \int_{1/\lambda}^{\infty} x \lambda e^{-\lambda x} dx =$$

$$\left\{ \begin{array}{l} u = x,\, du = dx \\ dv = \lambda e^{-\lambda x} dx,\, v = -e^{-\lambda x} \end{array} \right\} = e \left(-e^{-\lambda x} x \Big|_{1/\lambda}^{\infty} \right) + e \int_{1/\lambda}^{\infty} e^{-\lambda x} dx$$

$$= e \cdot e^{-1} \cdot \frac{1}{\lambda} + e \left(-\frac{1}{\lambda} e^{-\lambda x} \Bigg|_{1/\lambda}^{\infty} \right) = \frac{1}{\lambda} + \frac{1}{\lambda} = \frac{2}{\lambda}.$$

Ahora despejamos el parámetro en función del primer momento:

$$\mu_1 = \frac{2}{\lambda} \Leftrightarrow \lambda = \frac{2}{\mu_1}.$$

A partir de esa expresión teórica definimos el estimador de momentos del parámetro, reemplazando el momento poblacional $\mu_1 = \mathbb{E}(X)$ por el muestral $m_1 = \bar{X}_n$:

$$\hat{\lambda} = \frac{2}{\bar{X}_n}.$$

3. Siguin $X_1, \ldots, X_n$ variables aleatòries independents idènticament distribuïdes amb funció de densitat

$$f(x;\alpha) = \frac{1+\alpha x}{2}\mathbb{I}_{[-1,1]}(x), \text{ amb } \alpha \in [-1,1].$$

(a) Doneu l'estimador de α obtingut pel mètode dels moments (al qual anomenarem $\hat{\alpha}_m$).

(b) És $\hat{\alpha}_m$ un estimador no esbiaixat de α?.

(c) Per a $n = 100$ s'han observat els valors $x_1, \ldots, x_n$ procedents de les variables $X_1, \ldots, X_n$. D'aquestes dades se sap el següent:

$$\sum_{i=1}^n x_i = -21.44148, \sum_{i=1}^n x_i^2 = 33.56329, \sum_{i=1}^n x_i^3 = -12.76414, \sum_{i=1}^n x_i^4 = 19.74995.$$

Doneu el valor de l'estimador $\hat{\alpha}_m$ per aquestes dades i calculeu l'error estàndard de l'estimador.

Resultat: (a) $\hat{\alpha}_m = 3\bar{X}_n$. (b) Sí. (c) $\hat{\alpha}_m = -0.6432444$, $\text{st.er}(\hat{\alpha}_m) = 0.1622734$

Resolució:

(a)

$$\mathbb{E}(X) = \int_{-1}^{1} x\frac{1+\alpha x}{2}\,dx = \left(\frac{x^2}{4} + \frac{\alpha x^3}{6}\right)\Big|_{-1}^{1} = \frac{\alpha}{3}.$$

Per tant, $\hat{\alpha}_m = 3\bar{X}_n$.

(b) És no esbiaixat:

$$\mathbb{E}(\hat{\alpha}_m) = 3\mathbb{E}(\bar{X}_n) = 3\mathbb{E}(X) = 3\frac{\alpha}{3} = \alpha.$$

(c) Per aquestes dades

$$\hat{\alpha}_m = 3\bar{x}_n = \frac{3}{n}\sum_{i=1}^n x_i = \frac{3}{100}(-21.44148) = -0.6432444.$$

L'error estàndard de l'estimador és la seva desviació estàndard. Com que

$$\text{Var}(\hat{\alpha}_m) = \text{Var}(3\bar{X}_n),$$

l'error estàndard de $\hat{\alpha}_m$ és 3 vegades l'error estàndard de $\bar{X}_n$, que podem estimar per $S_x/\sqrt{n}$, on

$$S_x^2 = \frac{1}{n-1}\sum_{i=1}^n (x_i - \bar{x}_n)^2 = \frac{1}{n-1}\left(\sum_{i=1}^n x_i^2 - n\bar{x}_n^2\right) = \left(33.56329 - 100(-.2144148)^2\right)/99 = 0.292585$$

Aleshores, l'error estàndard de $\hat{\alpha}_m$ l'estimarem com

$$\text{st.er}(\hat{\alpha}_m) = 3\sqrt{\frac{0.292585}{100}} = 0.1622734$$

4. Considera una muestra aleatoria simple de la variable $X \sim U([0,b])$, $b > 0$. Da las expresiones de los estimadores de momentos i de máxima verosimilitud de b. Busca sus distribuciones en el muestreo (exactas o asintóticas), calcula su errores estándar, sus sesgos i sus errores cuadráticos medios. ¿Cuál de los dos estimadores de b es más eficiente?

Resultat: $\text{ECM}_b(\hat{b}_{mom}) = b^2/(12n)$. $\text{ECM}_b(\hat{b}_{ML}) = 2b^2/((n+1)(n+2))$.

Resolució:

Estimador de momentos:

$\mu = \mathbb{E}(X) = b/2 \Longrightarrow \hat{b}_{mom} = 2\bar{X}_n$.

Distribución en el muestreo: $\mathbb{E}(\hat{b}_{mom}) = 2(\bar{X}_n) = 2b/2 = b$.

$\text{Var}(\hat{b}_{mom}) = 4\text{Var}(\bar{X}_n) = 4b^2/(12n) = b^2/(3n)$, porque $\sigma^2 = \text{Var}(X) = b^2/12$.

Además, por el TLC si n es *grande*, $\bar{X}_n \approx N(\mu, \sigma^2/n)$. Entonces,

$$\hat{b}_{mom} \approx N\left(b, \frac{b^2}{3n}\right).$$

Error estándar: $\sqrt{\text{Var}(\hat{b}_{mom})} = b/\sqrt{3n}$.

Sesgo: $\mathbb{E}(\hat{b}_{mom}) - b = 0$.

ECM: Como el sesgo es 0, el ECM coincide con la varianza del estimador que vale $\text{Var}(\hat{b}_{mom}) = b^2/(3n)$.

Estimador de máxima verosimilitud:

$L(b; x_1, \ldots, x_n) = \prod_{i=1}^n \frac{1}{b} I_{[0,b]}(x_i) = \frac{1}{b^n} I_{[x_{(n)},\infty)}(b)$

$\Longrightarrow \hat{b}_{ML} = X_{(n)} = \max\{X_1, \ldots, X_n\}$.

Distribución en el muestreo: Sabemos que la función de distribución del máximo de n v.a.i.i.d. es $F_{X_{(n)}}(x) = F(x)^n$, donde $F(x)$ es la función de distribución común a $X_1, \ldots, X_n$. En el caso de la $U([0,b])$, $F(x) = x/b$ si $x \in [0,b]$ (vale 0 si $x < 0$ y vale 1 si $x \geq 1$). Así, $F_{X_{(n)}}(x) = x^n/b^n$, y la función de densidad correspondiente es

$$f_{X_{(n)}}(x) = \frac{nx^{n-1}}{b^n} I_{[0,b]}(x),$$

que es también la de $\hat{b}_{ML}$.

Calculemos $\mathbb{E}(X_{(n)})$ y $\mathbb{E}(X_{(n)}^2)$:

$$\mathbb{E}(X_{(n)}) = \frac{1}{b^n}\int_0^b nx^n dx = \frac{n}{n+1}b,$$

$$\mathbb{E}(X_{(n)}^2) = \frac{1}{b^n}\int_0^b nx^{n+1} dx = \frac{n}{n+2}b^2.$$

Así,

$$\mathbb{E}(\hat{b}_{ML}) = \frac{n}{n+1}b = \left(1 - \frac{1}{n+1}\right)b = b - \frac{b}{n+1},$$

$$\text{Var}(\hat{b}_{ML}) = \frac{n}{n+2}b^2 - \frac{n^2}{(n+1)^2}b^2 = \frac{n}{(n+2)(n+1)^2}b^2.$$

Error estándar: $\sqrt{\text{Var}(\hat{b}_{ML})}$.

Sesgo: $\text{Sesgo}_b(\hat{b}_{ML}) = \mathbb{E}(\hat{b}_{ML}) - b = -b/(n+1)$.

ECM:

$$\text{ECM}_b(\hat{b}_{ML}) = \text{Var}(\hat{b}_{ML}) + \text{Sesgo}_b^2(\hat{b}_{ML})$$

$$= \frac{n}{(n+2)(n+1)^2}b^2 + \frac{1}{(n+1)^2}b^2 = \frac{2}{(n+1)(n+2)}b^2.$$

Comparación de los dos estimadores: El ECM del estimador máximo verosímil es menor que el del estimador de momentos si y solo si

$$\frac{2}{(n+1)(n+2)} < \frac{1}{3n} \Leftrightarrow (n+1)(n+2) > 6n$$

$$\Leftrightarrow n^2 + 3n + 2 > 6n \Leftrightarrow n^2 - 3n + (3/2)^2 > 1/4$$

$$\Leftrightarrow (n - (3/2))^2 > 1/4 \Leftrightarrow n - (3/2) > 1/2 \Leftrightarrow n > 2.$$

Además es fácil ver que para $n = 1$ o para $n = 2$ los dos estimadores tienen el mismo error cuadrático medio.

En cuanto al comportamiento cuando n es grande, el ECM del estimador de momentos es $b^2/(3n)$, que tiende a 0 a la velocidad de $1/n$. Diremos que

$$\text{ECM}_b(\hat{b}_{mom}) = O\left(\frac{1}{n}\right).$$

Por su parte, el ECM del estimador máximo verosímil es $2b^2/((n+1)(n+2))$, es decir, tiende a 0 como $1/n^2$, es decir,

$$\text{ECM}_b(\hat{b}_{ML}) = O\left(\frac{1}{n^2}\right).$$

Concluimos que, en este caso, el estimador máximo verosímil es asintóticamente más eficiente que el estimador de momentos porque el ECM del primero tiende a 0 mucho más rápido que el del segundo.
En la práctica, el estimador máximo verosímil es mucho más eficiente que el estimador de momentos incluso para tamaños muestrales pequeños.

5. Siguin $X_1, \ldots, X_n$ variables aleatòries independents idènticament distribuïdes amb funció de densitat

$$f(x;\lambda) = \frac{1}{2}\lambda \exp(-\lambda|x|),\ x \in \mathbb{R}\, \lambda > 0.$$

Siguin $x_1, \ldots, x_n$ observacions d'aquests variables.

(a) Doneu l'estimador de λ obtingut pel mètode dels moments.

(b) Doneu l'estimador de λ obtingut pel mètode de la màxima versemblança.

Resolució:

(a) Per simetria de $f(x)$ al voltant de 0,

$$\mu_1 = \mathbb{E}(X) = \int_{-\infty}^{\infty} x\frac{1}{2}\lambda e^{-\lambda|x|}dx = 0.$$

Calculem doncs el moment de segon ordre:

$$\mu_2 = \mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2\frac{1}{2}\lambda e^{-\lambda|x|}dx = 2\int_{0}^{\infty} x^2\frac{1}{2}\lambda e^{-\lambda x}dx = \int_{0}^{\infty} y^2\lambda e^{-\lambda y}dy = \mathbb{E}(Y^2),$$

on $Y \sim \text{Exp}(\lambda)$. Com que

$$\mathbb{E}(Y^2) = \text{Var(Y)} + \mathbb{E}(X)^2 = \frac{1}{\lambda^2} + \frac{1}{\lambda^2} = \frac{2}{\lambda^2},$$

aleshores

$$\mu_2 = \frac{2}{\lambda^2} \Rightarrow \lambda = \left(\frac{2}{\mu_2}\right)^{1/2}.$$

Aleshores, anomenant $m_2 = (1/n)\sum_{i=1}^n x_i^2$, l'estimador de λ pel mètode dels moments és

$$\hat{\lambda}_m = \left(\frac{2}{m_2}\right)^{1/2} = \left(\frac{2n}{\sum_{i=1}^n x_i^2}\right)^{1/2}.$$

(b) La funció de versemblança és

$$L(\lambda; x_1, \ldots, x_n) = \prod_{i=1}^n f(x_i;\lambda) = \prod_{i=1}^n \frac{1}{2}\lambda e^{-\lambda|x_i|} = \frac{1}{2^n}\lambda^n e^{-\lambda\sum_{i=1}^n |x_i|}.$$

El seu logaritme és

$$\ell(\lambda; x_1, \ldots, x_n) = n\log(2) + n\log(\lambda) - \lambda\sum_{i=1}^n |x_i|.$$

Derivem respecte λ:

$$\ell'(\lambda) = \frac{n}{\lambda} - \sum_{i=1}^{n} |x_i|,\ \ell''(\lambda) = -\frac{n}{\lambda^2} < 0.$$

Resolem $\ell'(\lambda) = 0$:

$$\frac{n}{\lambda} - \sum_{i=1}^{n} |x_i| = 0 \Leftrightarrow \lambda = \frac{n}{\sum_{i=1}^{n} |x_i|}.$$

Com que la segona derivada de $\ell(\lambda) < 0$ per tot $\lambda > 0$, aquest valor que resol $\ell'(\lambda) = 0$ és el màxim de $\ell(\lambda)$. És a dir, l'estimador màxim versemblant de λ és

$$\hat{\lambda}_{MV} = \frac{n}{\sum_{i=1}^{n} |x_i|}.$$

6. Sea $X \sim \gamma(\alpha, b)$, $\alpha > 0$, $b > 0$, con densidad

$$f_X(x|\alpha, b) = \frac{1}{\Gamma(\alpha) b^\alpha} x^{\alpha-1} e^{-x/b}$$

para $x > 0$. Se dice que $Y = X^{-1}$ tiene distribución *gamma invertida* con parámetros (α, b), y se denota como $IG(\alpha, b)$.

(a) Prueba que la densidad de Y es

$$f_Y(y; \alpha, b) = \frac{1}{\Gamma(\alpha) b^\alpha} \frac{1}{y^{\alpha+1}} e^{-1/(by)} I_{(0,\infty)}(y).$$

(b) Calcula la esperanza y la varianza de Y.

Resolució:
$\mathbb{E}(Y) = 1/(b(\alpha-1))$, $\mathrm{Var}(Y) = 1/(b^2(\alpha-1)^2(\alpha-2))$.

Resolució:

(a) Se tiene que $Y = 1/X = g(X)$, con $g(x) = 1/x$ una biyección de $(0, \infty)$ en sí mismo. La función inversa de g es $h(y) = 1/y$. Para calcular la función de densidad de Y en un $y \in (0, \infty)$ aplicamos la fórmula del cambio de variable univariante:

$$f_Y(y; \alpha, b) = f_X(h(y); \alpha, b)|h'(y)|$$
$$= \frac{1}{\Gamma(\alpha) b^\alpha} \frac{1}{y^{\alpha-1}} e^{-1/(by)} \frac{1}{y^2} = \frac{1}{\Gamma(\alpha) b^\alpha} \frac{1}{y^{\alpha+1}} e^{-1/(by)}.$$

(b) Calculemos $\mathbb{E}(Y^r)$ para un $r < \alpha$ genérico:

$$\mathbb{E}(Y^r) = \mathbb{E}(X^{-r}) = \int_0^\infty \frac{1}{x^r} \frac{1}{\Gamma(\alpha) b^\alpha} x^{\alpha-1} e^{-x/b}$$
$$= \frac{\Gamma(\alpha - r)}{\Gamma(\alpha) b^r} \int_0^\infty \frac{1}{\Gamma(\alpha - r) b^{\alpha - r}} x^{(\alpha - r)-1} e^{-x/b} = \frac{\Gamma(\alpha - r)}{\Gamma(\alpha) b^r}.$$

Así,

$$\mathbb{E}(Y) = \frac{\Gamma(\alpha-1)}{\Gamma(\alpha) b} = \frac{1}{(\alpha-1)b},$$
$$\mathbb{E}(Y^2) = \frac{\Gamma(\alpha-2)}{\Gamma(\alpha) b^2} = \frac{1}{(\alpha-1)(\alpha-2)b^2},$$
$$\mathrm{Var}(Y) = \mathbb{E}(Y^2) - \mathbb{E}(Y)^2 = \frac{1}{(\alpha-1)(\alpha-2)b^2} - \frac{1}{(\alpha-1)^2 b^2}$$
$$= \frac{(\alpha-1) - (\alpha-2)}{(\alpha-1)^2(\alpha-2)b^2} = \frac{1}{(\alpha-1)^2(\alpha-2)b^2}.$$

7. Se observa una muestra de una variable aleatoria X exponencial de parámetro λ (con $\mathbb{E}(X) = 1/\lambda$).

   (a) Prueba que el método de momentos y el de máxima verosimilitud conducen al mismo estimador $\hat{\lambda}$ del parámetro λ.

   (b) Prueba que la distribución en el muestreo de $\hat{\lambda}$ es una Gamma invertida e indica cuáles son sus parámetros.

   (c) Calcula el sesgo y el error cuadrático medio de $\hat{\lambda}$ como estimador de λ.

Resultat: (b) $\hat{\lambda} \sim IG\left(\alpha = n, b = \frac{1}{n\lambda}\right)$. (c) $\text{Sesgo}_\lambda(\hat{\lambda}) = \lambda/(n-1)$, $\text{ECM}_\lambda(\hat{\lambda}) = \frac{n^2+n-2}{(n-1)^2(n-2)}\lambda^2$.

Resolució:

(a) Como $\mu = \mathbb{E}(X) = 1/\lambda$, el estimador de momentos es $\hat{\lambda} = 1/\bar{X}_n$. La función de verosimilitud es

$$L(\lambda; x_1, \ldots, x_n) = \lambda^n e^{-\lambda \sum_{i=1}^n x_i},$$

y su logaritmo

$$\ell(\lambda) \equiv \ell(\lambda; x_1, \ldots, x_n) = n \log(\lambda) - \lambda \sum_{i=1}^n x_i.$$

Derivando

$$\ell'(\lambda) = \frac{n}{\lambda} - \sum_{i=1}^n x_i,$$

e igualando a 0, $\ell'(\lambda) = 0$, se obtiene que el EMV de λ es también $\hat{\lambda} = 1/\bar{X}_n$ (este valor es el máximo del logaritmo de la verosimilitud porque $\ell''(\lambda) = -n/\lambda^2 < 0$ para todo λ).

(b) Observar que $X_i \sim \text{Exp}(\lambda) \equiv \text{Gamma}(\alpha = 1, \beta = \lambda)$. Por tanto $(1/n)X_i \sim \text{Exp}(n\lambda) \equiv \text{Gamma}(\alpha = 1, \beta = n\lambda)$. Entonces, aplicando que la suma de Gammas independientes con el mismo parámetro β es también Gamma (con el mismo β, y su parámetro α es la suma de los α's), se tiene que

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i = \sum_{i=1}^n \frac{1}{n} X_i \sim \text{Gamma}(\alpha = n, \beta = n\lambda).$$

Por tanto, $\hat{\lambda} = 1/\bar{X}_n$ es una Gamma invertida de parámetros $\alpha = n$ y $\beta = n\lambda$. Escribiremos que

$$\hat{\lambda} \sim IG(\alpha = n, \beta = n\lambda)$$

o, usando el parámetro *scale* $b = 1/\beta$, que

$$\hat{\lambda} \sim IG\left(\alpha = n, b = \frac{1}{n\lambda}\right).$$

(c) En uno de los problemas del tema anterior vimos que si $Y \sim IG(\alpha, b)$ entonces $\mathbb{E}(Y) = 1/(b(\alpha-1))$, $\text{Var}(Y) = 1/(b^2(\alpha-1)^2(\alpha-2))$. Entonces

$$\mathbb{E}(\hat{\lambda}) = \frac{1}{\frac{1}{n\lambda}(n-1)} = \frac{n}{n-1}\lambda = \left(1 + \frac{1}{n-1}\right)\lambda,$$

luego el sesgo de $\hat{\lambda}$ como estimador de λ es $\text{Sesgo}_\lambda(\hat{\lambda}) = \lambda/(n-1)$. Y su varianza es

$$\text{Var}(\hat{\lambda}) = \frac{1}{\frac{1}{n^2\lambda^2}(n-1)^2(n-2)} = \frac{n^2}{(n-1)^2(n-2)}\lambda^2.$$

Por tanto, el error cuadrático medio de $\hat{\lambda}$ como estimador de λ será

$$\text{ECM}_\lambda(\hat{\lambda}) = \text{Var}(\hat{\lambda}) + \text{Sesgo}^2_\lambda(\hat{\lambda})$$

$$= \frac{n^2}{(n-1)^2(n-2)}\lambda^2 + \frac{1}{(n-1)^2}\lambda^2 = \frac{n^2+n-2}{(n-1)^2(n-2)}\lambda^2.$$

8. Sean $X_1, \ldots, X_n$ v.a. independientes tales que $X_i \sim \text{Poisson}(i\theta)$, $\theta > 0$. Encuentra el estimador máximo verosímil de θ, prueba que es insesgado y calcula su varianza.

Resultat: $\hat{\theta} = \frac{1}{a_n} \sum_{i=1}^n X_i$ con $a_n = \sum_{i=1}^n i = \frac{n(n+1)}{2}$, $\text{Var}(\hat{\theta}) = \theta / a_n$.

Resolució:
Recordemos que la función de probabilidad de la Poisson de parámetro λ es

$$\Pr(X = x; \lambda) = e^{-\lambda} \frac{\lambda^x}{x!}.$$

Así la función de verosimilitud es

$$L(\theta; x_1, \ldots, x_n) = \prod_{i=1}^n \Pr(X_i = x_i; i\theta) = \prod_{i=1}^n e^{-i\theta} \frac{(i\theta)^{x_i}}{x_i!}$$

y su logaritmo

$$\ell(\theta) \equiv \ell(\theta; x_1, \ldots, x_n) = -\theta \sum_{i=1}^n i + \log(\theta) \sum_{i=1}^n x_i + \sum_{i=1}^n x_i \log(i) - \sum_{i=1}^n \log(x_i!).$$

Entonces, llamando $a_n = \sum_{i=1}^n i = n(n+1)/2$,

$$\ell'(\theta) = -a_n + \frac{\sum_{i=1}^n x_i}{\theta}.$$

Resolviendo la ecuación $\ell'(\theta) = 0$, obtenemos

$$\hat{\theta} = \frac{1}{a_n} \sum_{i=1}^n x_i.$$

Para confirmar que es un máximo calculamos la segunda derivada y vemos que es negativa para todo θ:

$$\ell''(\theta) = -\frac{\sum_{i=1}^n x_i}{\theta^2} < 0.$$

La esperanza de $\hat{\theta}$ es

$$\mathbb{E}(\hat{\theta}) = \frac{1}{a_n} \sum_{i=1}^n \mathbb{E}(X_i) = \frac{1}{a_n} \sum_{i=1}^n i\theta = \frac{\theta}{a_n} \sum_{i=1}^n i = \theta.$$

Su varianza es

$$\text{Var}(\hat{\theta}) = \frac{1}{a_n^2} \sum_{i=1}^n \text{Var}(X_i) = \frac{1}{a_n^2} \sum_{i=1}^n i\theta = \frac{\theta}{a_n^2} \sum_{i=1}^n i = \frac{\theta}{a_n}.$$

También podemos determinar la distribución en el muestreo de $\hat{\theta} = \left(\sum_{i=1}^n X_i\right) / a_n$. Dado que la suma de v.a. de Poisson independientes es también Poisson (con parámetro igual a la suma de los parámetros de las variables sumadas) se tiene que

$$\hat{\theta} = \frac{\sum_{i=1}^n X_i}{a_n} \sim \frac{1}{a_n} \text{Poisson}\left(\sum_{i=1}^n i\theta\right) \equiv \frac{1}{a_n} \text{Poisson}(a_n \theta).$$

De ahí podríamos haber obtenido también el valor de la esperanza y la varianza de $\hat{\theta}$.

9. Sean $X_1, \ldots, X_n$ v.a. independientes tales que $X_i \sim N(i\theta, 1)$, $\theta \in \mathbb{R}$. Encuentra el estimador máximo verosímil de θ, prueba que es insesgado y calcula su varianza.

Resolució:
$\hat{\theta} = \frac{1}{b_n} \sum_{i=1}^n iX_i$ con $b_n = \sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{6}$, $\text{Var}(\hat{\theta}) = 1/b_n$.

Resolució:
La función de densidad de X_i es

$$f_{X_i}(x_i) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x_i - i\theta)^2}$$

y su logaritmo

$$\log f_{X_i}(x_i) = -\frac{1}{2}\log(2\pi) - \frac{1}{2}(x_i - i\theta)^2 = -\frac{\log(2\pi)}{2} - \frac{1}{2}(x_i^2 - 2ix_i\theta + i^2\theta^2).$$

Así, el logaritmo de la función de verosimilitud es

$$\ell(\theta) \equiv \ell(\theta; x_1, \ldots, x_n) = \sum_{i=1}^n \log f_{X_i}(x_i)$$

$$= -\frac{n\log(2\pi)}{2} - \frac{1}{2}\sum_{i=1}^n (x_i^2 - 2ix_i\theta + i^2\theta^2).$$

Entonces,

$$\ell'(\theta) = -\frac{1}{2}\sum_{i=1}^n (-2ix_i + 2i^2\theta) = \sum_{i=1}^n ix_i - \theta\sum_{i=1}^n i^2.$$

Resolviendo la ecuación $\ell'(\theta) = 0$, obtenemos

$$\hat{\theta} = \frac{1}{b_n}\sum_{i=1}^n ix_i,$$

donde $b_n = \sum_{i=1}^n i^2 = n(n+1)(2n+1)/6$. Para confirmar que es un máximo calculamos la segunda derivada y vemos que es negativa para todo θ:

$$\ell''(\theta) = -b_n < 0.$$

La esperanza de $\hat{\theta}$ es

$$\mathbb{E}(\hat{\theta}) = \frac{1}{b_n}\sum_{i=1}^n i\mathbb{E}(X_i) = \frac{1}{b_n}\sum_{i=1}^n i^2\theta = \frac{\theta}{b_n}\sum_{i=1}^n i^2 = \theta.$$

Su varianza es

$$\mathrm{Var}(\hat{\theta}) = \frac{1}{b_n^2}\sum_{i=1}^n i\mathrm{Var}(X_i) = \frac{1}{b_n^2}\sum_{i=1}^n i^2 = \frac{1}{b_n}.$$

10. Sean $X_1, \ldots, X_n$ v.a. independientes tales que $X_i \sim \mathrm{Exp}(1/(i\theta))$, $E(X_i) = i\theta$, $\theta > 0$. Encuentra el estimador máximo verosímil de θ, prueba que es insesgado y calcula su varianza.

Resultat: $\hat{\theta} = \frac{1}{n}\sum_{i=1}^n \frac{X_i}{i}$, $\mathrm{Var}(\hat{\theta}) = \theta^2/n$.

Resolució:
La función de densidad de X_i es

$$f_{X_i}(x_i) = \frac{1}{i\theta} e^{-\frac{x_i}{i\theta}}$$

y su logaritmo

$$\log f_{X_i}(x_i) = -\log(i) - \log(\theta) - \frac{x_i}{i\theta}.$$

Así, el logaritmo de la función de verosimilitud es

$$\ell(\theta) \equiv \ell(\theta; x_1, \ldots, x_n) = \sum_{i=1}^n \log f_{X_i}(x_i)$$

$$= -\sum_{i=1}^{n} \log(i) - n\log(\theta) - \frac{1}{\theta}\sum_{i=1}^{n}\frac{x_i}{i}.$$

Entonces,

$$\ell'(\theta) = -\frac{n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n}\frac{x_i}{i}.$$

Resolviendo la ecuación $\ell'(\theta) = 0$, obtenemos

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n}\frac{x_i}{i}.$$

Para confirmar que es un máximo calculamos la segunda derivada y vemos que es negativa para todo $\theta = \hat{\theta}$:

$$\ell''(\theta) = \frac{n}{\theta^2} - \frac{2}{\theta^3}\sum_{i=1}^{n}\frac{x_i}{i}$$

y, por tanto,

$$\ell''(\hat{\theta}) = \frac{n}{\hat{\theta}^2} - \frac{2}{\hat{\theta}^3}n\hat{\theta} = -\frac{n}{\hat{\theta}^2} < 0.$$

La esperanza de $\hat{\theta}$ es

$$\mathbb{E}(\hat{\theta}) = \frac{1}{n}\sum_{i=1}^{n}\frac{\mathbb{E}(X_i)}{i} = \frac{1}{n}\sum_{i=1}^{n}\frac{i\theta}{i} = \frac{\theta}{n}\sum_{i=1}^{n}1 = \theta.$$

Su varianza es

$$\operatorname{Var}(\hat{\theta}) = \frac{1}{n^2}\sum_{i=1}^{n}\frac{\operatorname{Var}(X_i)}{i^2} = \frac{1}{n^2}\sum_{i=1}^{n}\frac{i^2\theta^2}{i^2} = \frac{\theta^2}{n^2}\sum_{i=1}^{n}1 = \frac{\theta^2}{n}.$$

11. **Hardy-Weinberg law.** (Problem 6.2.19 of Evans and Rosenthal, 2010).
The Hardy-Weinberg law in genetics says that the proportions of genotypes AA, Aa, and aa are θ^2, $2\theta(1-\theta)$, and $(1-\theta)^2$, respectively, where $\theta \in [0,1]$. Suppose that in a sample of n individuals from the population (small relative to the size of the population), we observe x_1 individuals of type AA, x_2 individuals of type Aa, and x_3 individuals of type aa.
   (a) What distribution do the counts (X_1, X_2, X_3) follow?
   (b) Record the likelihood function, the log-likelihood function, and the score function for θ.
   (c) Record the form of the MLE for θ.

Resultat:

Resolució:

12. Consideramos una muestra de X exponencial de parámetro λ (con $\mathbb{E}(X) = 1/\lambda$), como en el problema 7. Da la distribución asintótica de $\hat{\lambda}$ (el estimador de momentos y también el máximo verosímil de λ) usando la teoría asintótica del estimador máximo verosímil.

Resultat:

Resolució:
El estimador es $\hat{\lambda} = 1/\bar{X}_n$. La teoría asintótica del estimador máximo verosímil nos dice que

$$\sqrt{n}\left(\hat{\lambda}_n - \lambda\right) \underset{n\to\infty}{\longrightarrow} N\left(0, \frac{1}{I(\lambda)}\right) \text{ en distribución,}$$

donde $I(\lambda)$ es la Información de Fisher que sobre λ contiene el modelo (o la que contiene una observación de X):

$$I(\lambda) = \operatorname{Var}_\lambda\left(\frac{\partial \log(f(X;\lambda))}{\partial\lambda}\right) = -\mathbb{E}_\lambda\left(\frac{\partial^2 \log(f(X;\lambda))}{\partial\lambda^2}\right).$$

Calculemos esas cantidades:

$$f(x;\lambda)=\lambda e^{-\lambda x}\Rightarrow \log(f(x;\lambda))=\log(\lambda)-\lambda x$$

$$\Rightarrow \frac{\partial \log(f(x;\lambda))}{\partial\lambda}=\frac{1}{\lambda}-x\Rightarrow \frac{\partial^2 \log(f(x;\lambda))}{\partial\lambda^2}=-\frac{1}{\lambda^2}.$$

Observar que las dos formas alternativas de calcular la Información de Fisher dan lugar al mismo resultado:

$$\mathrm{Var}_\lambda\left(\frac{\partial \log(f(X;\lambda))}{\partial\lambda}\right)=\mathrm{Var}_\lambda\left(\frac{1}{\lambda}-X\right)=\mathrm{Var}_\lambda(X)=\frac{1}{\lambda^2},$$

$$-\mathbb{E}_\lambda\left(\frac{\partial^2 \log(f(X;\lambda))}{\partial\lambda^2}\right)=-\mathbb{E}_\lambda\left(-\frac{1}{\lambda^2}\right)=\frac{1}{\lambda^2}.$$

Así, $I(\lambda)=1/\lambda^2$ y, por tanto,

$$\sqrt{n}\left(\hat\lambda_n-\lambda\right)\underset{n\to\infty}{\longrightarrow} N\left(0,\frac{1}{I(\lambda)}=\lambda^2\right)\text{ en distribución.}$$

13. Sigui $X\sim$ Bin(n,p). Es defineix el *log-odds* com $\psi=\log(p/(1-p))$. Doneu l'estimador màxim versemblant de ψ i indiqueu quina és la seva distribució asimptòtica fent servir la Informació de Fisher.

Resultat:

Resolució:
Observar que $\psi=\tau(p)=\log(p/(1-p))$ y que $\tau(p)$ es una función biyectiva. Entonces, por el principio de invarianza, el estimador máximo verosímil de ψ es

$$\hat\psi_n=\tau(\hat p_n)=\log\left(\frac{\hat p_n}{1-\hat p_n}\right)$$

donde $\hat p_n=X/n$ es la proporción muestral de éxitos en los primeros n experimentos, que sabemos que es el estimador máximo verosímil de p. (Usamos el subíndice n porque vamos a hablar del comportamiento asintótico del estimador).
La teoría asintótica del estimador máximo verosímil nos dice que

$$\sqrt{n}\left(\hat\psi_n-\psi\right)=\sqrt{n}\left(\tau(\hat p_n)-\tau(p)\right)\underset{n\to\infty}{\longrightarrow} N\left(0,\frac{(\tau'(p))^2}{I(p)}\right)$$

en distribución, donde $I(p)$ es la Información de Fisher que sobre p contiene el modelo (o la que contiene una v.a. Y Bernouilli de parámetro p):

$$I(p)=\mathrm{Var}_p\left(\frac{\partial}{\partial p}\log(p^Y(1-p)^{(1-Y)})\right)$$

$$=\mathrm{Var}_p\left(\frac{\partial}{\partial p}\{Y\log(p)+(1-Y)\log(1-p)\}\right)$$

$$=\mathrm{Var}_p\left(\frac{Y}{p}-\frac{1-Y}{1-p}\right)=\mathrm{Var}_p\left(\frac{Y-1}{p(1-p)}\right)$$

$$=\frac{p(1-p)}{p^2(1-p)^2}=\frac{1}{p(1-p)}.$$

También podríamos haber calculado $I(p)$ así:

$$I(p)=-\mathbb{E}_p\left(\frac{\partial^2}{\partial p^2}\log(p^Y(1-p)^{(1-Y)})\right)$$

$$=-\mathbb{E}_p\left(\frac{\partial}{\partial p}\left\{\frac{Y}{p}-\frac{1-Y}{1-p}\right\}\right)=-\mathbb{E}_p\left(-\frac{Y}{p^2}-\frac{1-Y}{(1-p)^2}\right)$$

$$= \frac{1}{p} + \frac{1}{1-p} = \frac{1}{p(1-p)}.$$

Calculemos ahora $\tau'(p)$:

$$\tau'(p) = \frac{d}{dp} \log\left(\frac{p}{1-p}\right) = \frac{1-p}{p} \frac{(1-p) - p(-1)}{(1-p)^2} = \frac{1}{p(1-p)}.$$

Así concluimos que

$$\sqrt{n}\left(\hat{\psi}_n - \psi\right) \underset{n\to\infty}{\longrightarrow} N\left(0, \frac{\frac{1}{p^2(1-p)^2}}{\frac{1}{p(1-p)}} = \frac{1}{p(1-p)}\right).$$