Ayudantía 9 - Mat044

17 de noviembre de 2021

Distribución Normal

Considerando una muestra aleatoria de tamaño n de X_i donde $X_i \sim^{iid} N(\mu, \sigma^2)$, como son iid se puede escribir también en la forma multivaraida $X = (X_1, \ldots, X_n) \sim N_n(\mu, \sigma^2 I_n)$.

Entonces, al considerar la suma de los X_i se tiene que

$$X = \sum X_i \Rightarrow E(X) = E\left(\sum X_i\right) = \sum E(X_i) = n\mu$$

Luego, al considerar $X = \frac{1}{n} \sum X_i$ se tiene que

$$E(X) = \mu$$

Finalmente, si se tiene una muestra de tamaño n de una normal (iid) se tiene que el promedio aritmético de los datos es una v.a. y además tiene esperanza μ.

Si $E(\bar{x}) = \mu$, ¿Qué ocurre con su varianza?

Como la muestra es independiente, $X_i \perp X_j \ \forall i \neq j$ entonces se tiene que $V(X_i + X_j) = V(X_i) + V(X_j)$. Además, la varianza es función cuadrática, por lo tanto

$$V(\alpha x) = \alpha^2 V(x)$$

De este modo,

$$V(\bar{x}) = V\left(\frac{1}{n}_i\right) = \frac{1}{n^2}\sum V(X_i) = \frac{1}{n^2}\sum \sigma^2 = \frac{\sigma^2}{n}$$

Por lo tanto, se tiene que si $X_i \sim^{iid} N(\mu, \sigma^2)$ es una muestra aleatoria, entonces $\bar{X}$ es una variable aleatoria y además tenemos que

$$\bar{X} \sim N(\mu, \sigma^2/n).$$

Entonces se tienen las siguientes características:

- **Insesgado :** La esperanza del estimador es el parámetro a estimar: $E(\bar{x}) = \mu$
- **Consistente :** $\lim_{n\to\infty} \bar{X} = \mu \, [\forall \epsilon > 0 \ \ \lim_{n\to\infty} P(|\bar{x}_n . \mu| > \epsilon = 0)]$

Esto último quiere decir, que a medida que el tamaño de la muestra n crece $\bar{x}$ 'converge' a un valor determinístico , '$\lim_{n\to\infty} \bar{x}_n \sim N(\mu, 0)$'.

Estimación: método de momentos

Para una muestra $X_1, \ldots, X_n$ que sigue un modelo estadístico $\{P_\theta^n : \theta \in \Theta\}$ se tiene el estimador de momentos como:

$$m_k = \frac{1}{n} \sum x_i^k$$

donde $m_k = \hat{\mu}_k$ k-ésimo momento empírico.

Considerando la siguiente muestra aleatoria:

i	X_i	i	X_i
1	172,45	11	141,14
2	141,93	12	71,19
3	146,82	13	83,39
4	172,66	14	128,33
5	96,94	15	90,82
6	100,20	16	94,37
7	161,51	17	124,05
8	104,76	18	120,68
9	128,06	19	110,60
10	156,38	20	125,34

Si se desea obtener los parámetros de interés de la media y desviación estándar μ y σ, si los datos provienen de una normal entonces:

$$\mu = \int x f(x;\theta) dx$$

$$\sigma^2 = \int (x-\mu)^2 f(x;\theta) dx$$

Entonces, se tiene el estimador de momentos para μ (es el primer momento)

$$m_1 = \frac{1}{n} \sum x_i = 114{,}37$$

luego, para obtener el estimador de momentos para :

$$\sigma^2 = V(X) = E(X^2) - E(X)^2 = \mu_2 - \mu_1^2$$

$$\sigma = \sqrt{\mu_2 - \mu_1^2}$$

$$\hat{\sigma}_{MM} = \sqrt{m_2 - m_1^2}$$

$$m_2 - m_1^2 = \frac{1}{n}\sum x_i^2 - \left(\frac{1}{n}\sum x_i\right)^2 = \frac{1}{n}\left(\sum x_i^2 - \frac{1}{n}(n\bar{x})^2\right)$$
$$= \frac{1}{n}\left(\sum x_i^2 - n\bar{x}^2\right) = \frac{1}{n}\sum (x_i - \bar{x})^2$$

entonces, se tiene que

$$\hat{\sigma}_{MM} = \sqrt{\frac{1}{n}\sum (x_i - \bar{x})^2}$$

donde para la muestra se tiene que

$$\sum (x_i - \bar{x})^2 = 22734,86$$

por lo tanto

$$\hat{\sigma}_{MM} = 33,72$$

Estimación máximo verosimil

Considerando una muestra aleatoria x_i iid con función de probabilidad $f(\cdot)$ la cual pertenece a un modelo estadístico $P_\theta = \{f(\cdot|\theta),\ \theta \in \Theta\}$ de manera que $f(\cdot) = f(\cdot|\theta_0)$ donde θ_0 es el parámetro real de donde provienen los datos.

Se desea inferir el parámetro $\hat{\theta}$ tal que sea lo más cercano posible al valor real θ_0 a través de la muestra x_i.

Se tiene la densidad conjunto de la uestra:

$$L(\theta_0|x_1,\ldots,x_n) = f(x_1,\ldots,x_n|\theta_0) = \prod_i f(x_i|\theta_0)$$

pues la muestra es independiente, donde $L(_0|x_i)$ se llama función de verosimilitud (likelihood)

Considernado la función de verosimilitud:

$$L(\theta|x_i) = \prod_i f(x_i|\theta)$$

se busca el parámetro $\hat{\theta}$ tal que maximice la probabilidad de obtener dicha muestra, entonces se tiene el estimador máximo verosímil (MLE):

$$\hat{\theta}_{MLE} \in \arg\max_{\theta \in \Theta} L(\theta|x_i)$$

Como $L(\theta)$ es una productoria, se tiene $ln(L) = \sum ln(f(\theta|x_i))$, como ln es una función monótona creciente se tiene que

$$\hat{\theta}_{MLE} \in \arg\max_{\theta \in \Theta} L(\theta|x_i) \iff \hat{\theta}_{MLE} \in \arg\max_{\theta \in \Theta} ln(L(\theta|x_i))$$

por lo que es más fácil trabajar con la log-verosimilitud:

$$\ell(\theta) = ln(L(\theta))$$

entonces el método se reduce a encontrar el parámetro $\hat{\theta}$ tal que maximice la log-verosimilitud (o verosimilitud a conveniencia).

Caso Normal

Considerando una muestra aleatoria desde una $N(\mu, \sigma^2)$, la verosimilitud:

$$L(\mu, \sigma^2; x) = (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2\right\}$$

considerando la log-verosimilitud

$$\ell(\mu, \sigma^2) = -\frac{n}{2} ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2$$

entonces, para maximizar verosimilitud se tiene que :

$$\begin{aligned} \frac{\partial \ell}{\partial \mu}(\mu, \sigma^2) = & \quad \tfrac{1}{\sigma^2} \textstyle\sum_{i=1}^n (x_i - \mu) & = 0 \\ \frac{\partial \ell}{\partial \sigma^2}(\mu, \sigma^2) = & \quad -\tfrac{n}{2\sigma^2} + \tfrac{1}{2\sigma^4} \textstyle\sum_{i=1}^n (x_i - \mu)^2 & = 0 \end{aligned}$$

y se tiene que el máximo se alcanza en:

$$\hat{\mu}_{ML} = \bar{x}, \quad \hat{\sigma}^2_{ML} = \frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{\mu}_{ML})^2$$

Volviendo al caso inicial con los estimadores de momentos, se tiene el estimador ML pra σ^2 pero por una propiedad de invarianza, se tiene que el MLE para σ sería $\sqrt{\hat{\sigma}^2_{ML}}$, entonces, para una muestra aleatoria que proviene de una distribución normal, el estimador ML coincide con el estimador de momentos.

Considerando $X_1, \ldots, X_n$ una muestra de v.a. IID desde una $U[0,\theta]$. Esto es,

$$f(x;\theta) = \begin{cases} \frac{1}{\theta}, & 0 \leq x \leq \theta \\ 0, & \text{en otro caso} \end{cases} = \frac{1}{\theta} I_{[0,\theta]}(x_i) =, \quad \theta > 0$$

de esta forma

$$L(\theta;x) = \frac{1}{\theta^n} \prod_{i=1}^{n} I_{[0,\theta]}(x_i) = \frac{1}{\theta^n} \prod_{i=1}^{n} I_{[x_i,\infty)}(\theta)$$

entonces, queremos encontrar el máximo de la verosimilitud

$$\prod_{i=1}^{n} I_{[x_i,\infty)}(\theta) = 1 \iff I_{[x_i,\infty)}(\theta) = 1 \ \forall i \iff x_i \leq \theta \forall i$$

esto es, $\max_i\{x_i\} \leq \theta \iff I_{[x_i,\infty)}(\theta) = 1$. Luego, se tiene que

$$L(\theta;x) = \frac{1}{\theta^n} I_{[\max_i x_i,\infty)}(\theta)$$

entonces

$$\hat{\theta}_{ML} = \max_i\{x_i\}$$

Si (2.85, 1.51, 0.69, 0.57, 2.29) es una muestra de una $U[0,\theta]$ entonces $\hat{\theta}_{ML} = 2{,}85$