STAT 400 Review (5)

p.m.f. or p.d.f. $f(x;\theta)$, $\theta \in \Omega$. Ω – parameter space.

1. Suppose $\Omega = \{1, 2, 3\}$ and the p.m.f. $f(x;\theta)$ is

$\theta = 1$: $f(1;1) = 0.6$, $f(2;1) = 0.1$, $f(3;1) = 0.1$, $f(4;1) = 0.2$.

$\theta = 2$: $f(1;2) = 0.2$, $f(2;2) = 0.3$, $f(3;2) = 0.3$, $f(4;2) = 0.2$.

$\theta = 3$: $f(1;3) = 0.3$, $f(2;3) = 0.4$, $f(3;3) = 0.2$, $f(4;3) = 0.1$.

What is the maximum likelihood estimate of θ based on only one observation of X if …

a) $X = 1$;

$$\begin{aligned} &f(1;1) = 0.6 \Leftarrow \\ &f(1;2) = 0.2 \qquad \Rightarrow \qquad \hat{\theta} = \mathbf{1}. \\ &f(1;3) = 0.3 \end{aligned}$$

b) $X = 2$;

$$\begin{aligned} &f(2;1) = 0.1 \\ &f(2;2) = 0.3 \qquad \Rightarrow \qquad \hat{\theta} = \mathbf{3}. \\ &f(2;3) = 0.4 \Leftarrow \end{aligned}$$

c) $X = 3$;

$$\begin{aligned} &f(3;1) = 0.1 \\ &f(3;2) = 0.3 \Leftarrow \qquad \Rightarrow \qquad \hat{\theta} = \mathbf{2}. \\ &f(3;3) = 0.2 \end{aligned}$$

d) $X = 4$.

$$\begin{aligned} &f(4;1) = 0.2 \Leftarrow \\ &f(4;2) = 0.2 \Leftarrow \qquad \Rightarrow \qquad \hat{\theta} = \mathbf{1 \text{ or } 2}. \\ &f(4;3) = 0.1 \end{aligned}$$

(maximum likelihood estimate may not be unique)

Likelihood function:

$$\mathrm{L}(\theta) = \mathrm{L}(\theta; x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} f(x_i; \theta) = f(x_1; \theta) \cdot \ldots \cdot f(x_n; \theta)$$

It is often easier to consider $\ln \mathrm{L}(\theta) = \sum\limits_{i=1}^{n} \ln f(x_i; \theta)$.

Maximum Likelihood Estimator: $\hat{\theta} = \arg\max \mathrm{L}(\theta) = \arg\max \ln \mathrm{L}(\theta)$.

Method of Moments:

$\mathrm{E}(\mathrm{X}) = \mathrm{g}(\theta)$. Set $\overline{\mathrm{X}} = \mathrm{g}(\tilde{\theta})$. Solve for $\tilde{\theta}$.

0. Consider a single observation X of a Binomial random variable with n trials and probability of "success" p. That is,

$$\mathrm{P}(\mathrm{X} = k) = {}_nC_k\, p^{k} (1-p)^{n-k}, \qquad k = 0, 1, \ldots, n.$$

a) Obtain the method of moments estimator of p, $\tilde{p}$.

Binomial: $\mathrm{E}(\mathrm{X}) = n\,p$

$\mathrm{X} = n\,\tilde{p} \qquad \Rightarrow \qquad \tilde{p} = \dfrac{\mathrm{X}}{n}$.

b) Obtain the maximum likelihood estimator of p, $\hat{p}$.

$$\mathrm{L}(p) = {}_nC_{\mathrm{X}}\, p^{\mathrm{X}} (1-p)^{n-\mathrm{X}}$$

$$\ln \mathrm{L}(p) = \ln {}_nC_{\mathrm{X}} + \mathrm{X} \ln p + (n - \mathrm{X}) \ln(1-p)$$

$$\frac{d}{dp} \ln \mathrm{L}(p) = \frac{\mathrm{X}}{p} - \frac{n-\mathrm{X}}{1-p} = \frac{\mathrm{X} - \mathrm{X}p - np + \mathrm{X}p}{p(1-p)} = \frac{\mathrm{X} - np}{p(1-p)}$$

$$\frac{d}{dp} \ln \mathrm{L}(\hat{p}) = 0 \qquad \Rightarrow \qquad \hat{p} = \frac{\mathrm{X}}{n}.$$

2. Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from a Poisson distribution with mean λ, $\lambda > 0$. That is,

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad k = 0, 1, 2, 3, \ldots.$$

a) Obtain the method of moments estimator of λ, $\tilde{\lambda}$.

$$E(X) = \lambda \qquad \Rightarrow \qquad \tilde{\lambda} = \overline{X}$$

b) Obtain the maximum likelihood estimator of λ, $\hat{\lambda}$.

$$L(\lambda) = \prod_{i=1}^{n} f(X_i; \lambda) = \prod_{i=1}^{n} \left(\frac{\lambda^{X_i} e^{-\lambda}}{X_i!} \right).$$

$$\ln L(\lambda) = \left(\sum_{i=1}^{n} X_i \right) \cdot \ln \lambda - n\lambda - \sum_{i=1}^{n} \ln(X_i!).$$

$$\frac{d \ln L(\lambda)}{d\lambda} = \frac{1}{\lambda} \cdot \left(\sum_{i=1}^{n} X_i \right) - n = 0. \qquad \Rightarrow \qquad \hat{\lambda} = \frac{1}{n} \sum_{i=1}^{n} X_i = \overline{X}.$$

Let $\hat{\theta}$ be the maximum likelihood estimate (m.l.e.) of θ. Then the m.l.e. of any function $h(\theta)$ is $h(\hat{\theta})$. (The Invariance Principle)

c) Obtain the maximum likelihood estimator of $P(X = 2)$.

$$P(X = 2) = h(\lambda) = \frac{\lambda^2 e^{-\lambda}}{2!} \qquad \hat{\lambda} = \overline{X} \qquad h(\hat{\lambda}) = \frac{\overline{X}^2 e^{-\overline{X}}}{2!}.$$

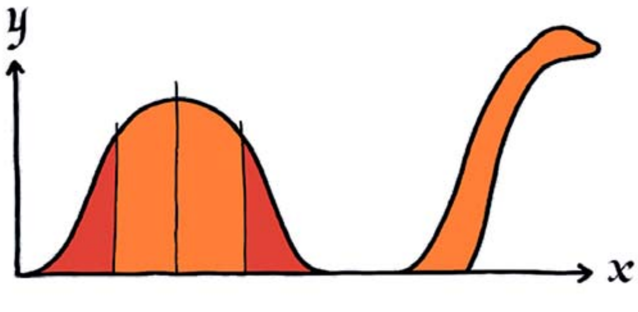
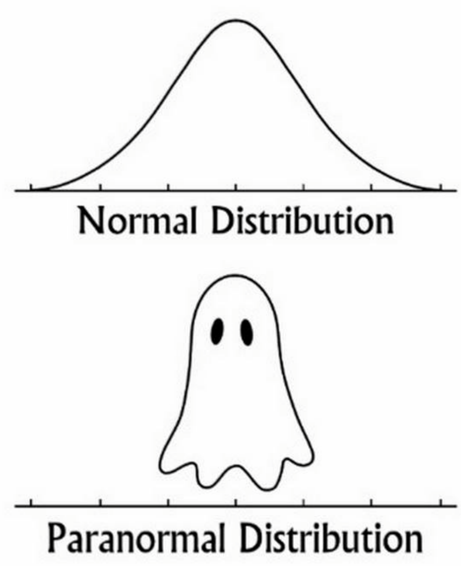


fig 1.0 The Extended Bell Curve.



3. Let $X_1, X_2, \ldots, X_n$ be a random sample from $N(\mu, \sigma^2)$, μ unknown.

Show that $\hat{\mu} = \overline{X}$ is the MLE for μ.

$$L(\mu; x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{1}{2\sigma^2}(x_i - \mu)^2 \right\}.$$

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

$$\frac{d}{d\mu}\ln L(\mu) = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu) = 0.$$

$$\sum_{i=1}^{n}(x_i - \mu) = \sum_{i=1}^{n} x_i - n\mu = 0.$$

$$\Rightarrow \qquad \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}.$$

4. Let $X_1, X_2, \ldots, X_n$ be a random sample from $N(\mu, \sigma^2)$, μ known, σ unknown. Show that $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$ is the MLE for σ^2.

$$L(\sigma^2; x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x_i - \mu)^2\right\}.$$

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

$$\frac{d}{d\sigma}\ln L(\sigma) = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i - \mu)^2 = 0.$$

$$\Rightarrow \qquad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2.$$

OR

$$L(\sigma^2; x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x_i - \mu)^2\right\}$$

$$= \left(\frac{1}{\sqrt{2\pi\theta}}\right)^n \exp\left\{-\frac{1}{2\theta}\sum_{i=1}^{n}(x_i - \mu)^2\right\}. \qquad \theta = \sigma^2$$

$$\ln L(\theta) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\theta) - \frac{1}{2\theta}\sum_{i=1}^{n}(x_i - \mu)^2.$$

$$\frac{d}{d\theta}\ln L(\theta) = -\frac{n}{2\theta} + \frac{1}{2\theta^2}\sum_{i=1}^{n}(x_i - \mu)^2 = 0.$$

$$\Rightarrow \qquad \hat{\theta} = \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2.$$

5. Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from $N(\theta_1, \theta_2)$, where $\Omega = \{(\theta_1, \theta_2) : -\infty < \theta_1 < \infty,\ 0 < \theta_2 < \infty\}$. That is, here we let $\theta_1 = \mu$ and $\theta_2 = \sigma^2$.

a) Obtain the maximum likelihood estimator of θ_1, $\hat{\theta}_1$, and of θ_2, $\hat{\theta}_2$.

$$L(\theta_1, \theta_2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\theta_2}} \exp\left[-\frac{(X_i - \theta_1)^2}{2\theta_2}\right]$$

$$= \left(\frac{1}{\sqrt{2\pi\theta_2}}\right)^n \exp\left[\frac{-\sum_{i=1}^{n}(X_i - \theta_1)^2}{2\theta_2}\right], \qquad (\theta_1, \theta_2) \in \Omega.$$

$$\ln L(\theta_1, \theta_2) = -\frac{n}{2}\ln(2\pi\theta_2) - \frac{\sum_{i=1}^{n}(X_i - \theta_1)^2}{2\theta_2}.$$

The partial derivatives with respect to θ_1 and θ_2 are

$$\frac{\partial(\ln L)}{\partial\theta_1} = \frac{1}{\theta_2}\sum_{i=1}^{n}(X_i - \theta_1)$$

and

$$\frac{\partial(\ln L)}{\partial\theta_2} = -\frac{n}{2\theta_2} + \frac{1}{2\theta_2^2}\sum_{i=1}^{n}(X_i - \theta_1)^2.$$

The equation $\partial(\ln L)/\partial\theta_1 = 0$ has the solution $\theta_1 = \overline{X}$.

Setting $\partial(\ln L)/\partial\theta_2 = 0$ and replacing θ_1 by $\overline{X}$ yields

$$\theta_2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2.$$

Therefore, the maximum likelihood estimators of $\mu = \theta_1$ and $\sigma^2 = \theta_2$ are

$$\hat{\theta}_1 = \overline{X} \quad \text{and} \quad \hat{\theta}_2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2.$$

b) Obtain the method of moments estimator of θ_1, $\tilde{\theta}_1$, and of θ_2, $\tilde{\theta}_2$.

$E[X] = \mu = \theta_1$.

$E[X^2] = \operatorname{Var}[X] + E[X]^2 = \sigma^2 + \mu^2 = \theta_2 + \theta_1^2$.

Thus, $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = \tilde{\theta}_1$, $\quad \overline{X^2} = \frac{1}{n}\sum_{i=1}^{n} X_i^2 = \tilde{\theta}_2 + \tilde{\theta}_1^2$.

Therefore, $\tilde{\theta}_1 = \overline{X}$ and $\tilde{\theta}_2 = \overline{X^2} - \left(\overline{X}\right)^2$.