1. Exact Methods

2. Asymptotic methods

approximate

Example. $X_1, \cdots, X_n \overset{iid}{\sim} \text{Exp}(\theta)$. $E(X) = \theta$

$$\hat{\theta} = \frac{T}{n} \qquad T = \sum_{i=1}^{n} X_i$$

$100 \times (1-\alpha)\%$ CI of θ.

- Exact $100 \times (1-\alpha)\%$ CI of θ $[\hat{\theta}_L, \hat{\theta}_U]$ $\quad P(\hat{\theta}_L \le \theta \le \hat{\theta}_U) = 1-\alpha$

$$T \sim \text{Gamma}(n, \theta)$$

$$\frac{2T}{\theta} \sim \text{Gamma}(n, 2) = \chi^2_{2n}$$

$$\chi^2_{2n, 1-\frac{\alpha}{2}} \le \frac{2T}{\theta} \le \chi^2_{2n, \frac{\alpha}{2}}$$

$$\left[\frac{2n\hat{\theta}}{\chi^2_{2n, \frac{\alpha}{2}}}, \frac{2n\hat{\theta}}{\chi^2_{2n, 1-\frac{\alpha}{2}}} \right]$$

- Approximated $100 \times (1-\alpha)\%$ CI for θ:

$$\hat{\theta} \sim AN\left(\theta, \frac{\theta^2}{n}\right)$$

CLT: $\dfrac{\sqrt{n}(\hat{\theta} - \theta)}{\theta} \xrightarrow{d} N(0,1)$

$$-z_{\alpha/2} \le \frac{\sqrt{n}(\hat{\theta} - \theta)}{\theta} \le z_{\alpha/2}$$

$$\left[\frac{\sqrt{n}\hat{\theta}}{\sqrt{n} + z_{\alpha/2}}, \frac{\sqrt{n}\hat{\theta}}{\sqrt{n} - z_{\alpha/2}} \right]$$

Data $\underset{\sim}{Y} = (Y_1, \cdots, Y_n)^T$ $\quad f(\underset{\sim}{y}; \underset{\sim}{\theta})$ is joint density of $\underset{\sim}{Y}$

$$\underset{\sim}{\theta} = (\theta_1, \cdots, \theta_b)^T$$

Likelihood function is just $f(\underset{\sim}{Y}; \underset{\sim}{\theta}) = L(\underset{\sim}{\theta} \mid \underset{\sim}{Y})$

iid : $L(\underset{\sim}{\theta} \mid \underset{\sim}{Y}) = \prod_{i=1}^{n} f(Y_i; \underset{\sim}{\theta})$

independent $\quad Y_i \sim f_i(y; \theta)$

$$L(\underset{\sim}{\theta} \mid \underset{\sim}{Y}) = \prod_{i=1}^{n} f_i(Y_i; \underset{\sim}{\theta})$$

Important, In all situations, the likelihood is the joint density of the observed data to be analyzed

1. Interpretion
2. Is the density appropiate for data
3. obersved data?

(Interpretation) $\quad Y_i, i=1,\cdots,n \overset{iid}{\sim}$ Discrete r.v. with Prob. mass. funcion $f(y; \underset{\sim}{\theta})$ $\; // \;$ $Pr(Y = y; \underset{\sim}{\theta})$

$$L(\underset{\sim}{\theta} \mid \underset{\sim}{Y}) = \prod_{i=1}^{n} f(Y_i; \theta) = \prod_{i=1}^{n} Pr(Y_i^* = Y_i; \theta)$$

where $Y_i^*, i=1,\cdots,n$ are iid from $f(y;\theta)$ and are mutually independent of $Y_1 \cdots Y_n$

continuous

$$f(Y_i, \theta) = \lim_{h \to 0^+} \frac{1}{2h} \left(F(Y_i + h; \theta) - F(Y_i - h; \theta) \right)$$

$$= \lim_{h \to 0^+} \frac{Pr(Y_i - h \le Y_i^* \le Y_i + h)}{2h}$$

$$L(\underset{\sim}{\theta} \mid \underset{\sim}{Y}) = \lim_{h \to 0^+} \left(\frac{1}{2h}\right)^n \prod_{i=1}^{n} Pr(Y_i - h \le Y_i^* \le Y_i + h)$$

(Is the density appropriate?)

Example. 2.1 $\quad$ counts of movements in five-second intervals of one fetal Lamb.

No of Movements.	0	1	2	3	4	5	6	7	
counts	182	41	12	2	2	0	0	1	240

Starting with Poisson $f(y;\theta) = \dfrac{\theta^y e^{-\theta}}{y!}$ $y = 0, 1, 2, \cdots$.

$$\prod_{i=1}^{n} \frac{\theta^{y_i} e^{-\theta}}{y!} \overset{MLE}{\Rightarrow} \hat{\theta} = \bar{Y}_n \text{ (sample mean)}$$

$X_1 \cdots, X_n \overset{iid}{\sim} f_X(x; \underset{\sim}{\theta})$

$$\underset{\sim}{\hat{\theta}}^{(x)}_{MLE} = \underset{\underset{\sim}{\theta}}{\operatorname{argmax}}\ L(\underset{\sim}{\theta} | \underset{\sim}{X})$$

$$L(\underset{\sim}{\theta} | \underset{\sim}{X}) = \prod_{i=1}^{n} f_X(x; \underset{\sim}{\theta})$$

Suppose. g is a one-to-one transformation. $Y_i = g(X_i)$

$$\underset{\sim}{\hat{\theta}}^{(y)}_{MLE} = \underset{\underset{\sim}{\theta}}{\operatorname{argmax}}\ L(\underset{\sim}{\theta} | \underset{\sim}{Y})$$

$$f_y(y; \underset{\sim}{\theta}) = f_X(g^{-1}(y); \theta) \times |g^{-1}(y)'|$$

$$\underline{\hat{\theta} \quad L(\underset{\sim}{\theta} | \underset{\sim}{Y})} = \prod_{i=1}^{n} f_X(g^{-1}(Y_i); \theta) \times |g^{-1}(y_i)|'$$

$$= L(\theta | \underset{\sim}{X}) \times \text{constant}$$

$$\hat{\theta}^{(y)}_{MLE} = \hat{\theta}(Y_1, \cdots, Y_n)$$

$$\shortparallel$$

$$\hat{\theta}^{(x)}_{MLE} = \hat{\theta}(X_1, \cdots, X_n)$$

$X_1, \cdots, X_n \overset{iid}{\sim} N(\mu, 1)$ $\quad \hat{\theta}^{(x)}_M = \bar{X}_n = \dfrac{\Sigma X_i}{n}$

$$\shortparallel$$

$Y_i = \exp(X_i)$ $\quad \hat{\theta}^{(y)}_{MLE} = \dfrac{\Sigma \ln(Y_i)}{n}$

Empirical Estimator of F

suppose $Y_1, \dots, Y_n \overset{iid}{\sim} f(y)$ with cdf $F(y)$

- nonnegative, non-decreasing, right-continuous. $\lim_{y\to-\infty} F(y)=0$, $\lim_{y\to+\infty} F(y)=1$.

$$P(Y=Y_i) \approx \{\underbrace{F(Y_i+h)-F(Y_i-h)}_{P_{ih}}\}$$

$$L(F|\underset{\sim}{Y}) \approx \prod_{i=1}^{n} P_{ih}$$

$0<P_{ih}<1$ $\qquad \sum_{i=1}^{n} P_{ih}=1$

Estimate $P_{1h}, \dots, P_{nh}$, by maximizing $\prod_{i=1}^{n} P_{ih}$ w.r.t. $0<P_{ih}<1$, $\sum_{i=1}^{n} P_{ih}=1$

method of Lagrange Multipliers. find the stationary points of

$$g(P_{1h}, \dots, P_{nh}, \lambda) = \sum_{i=1}^{n} \log(P_{ih}) + \lambda\left(\sum_{i=1}^{n} P_{ih} - 1\right)$$

$$\frac{\partial g}{\partial P_{ih}} = \frac{1}{P_{ih}} + \lambda = 0 \qquad i=1,\dots,n$$

$$P_{ih} = -\lambda^{-1}$$

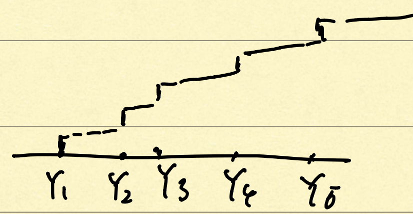
$$\frac{\partial g}{\partial \lambda} = \sum_{i=1}^{n} P_{ih} - 1 = 0$$

$$-\frac{n}{\lambda} - 1 = 0 \Rightarrow \lambda = -n$$

$$P_{ih} = \frac{1}{n}$$

$$\hat{P}_{ih} = \frac{1}{n} \quad \text{for } i=1,\dots,n$$

$$= F(Y_i+h) - F(Y_i-h)$$



$$\hat{F}_{MLE}(y) = \frac{1}{n}\sum_{i=1}^{n} I(Y_i \le y)$$

- Censoring. $X_1, \cdots, X_n \sim f(x;\underset{\sim}{\theta})$

$Y_i = \min(X_i, R_i)$

$\delta_i = I(X_i \le R_i) \begin{cases} = 1, & X_i \text{ before } R_i \\ = 0 & \text{censored. } X_i > R_i \end{cases}$

Example. pieces of equipment that are started at different times and late regularly checked for failure.

Y:	2	72	51	60	33	27	14	24	4	21
δ:	1	0	1	0	1	1	1	1	1	0

suppose. $\cancel{X_1 \sim X_{10}} \overset{iid}{\sim} \exp(\theta)$ How to estimate θ?

(Y_i, δ_i)

$P(Y_i \le y, \delta_i = 1) = P(Y_i \le y, Y_i = X_i) = P(X_i \le y)$

If $\delta_i = 1$. $Y_i \sim f_X(y;\theta)$

If $\delta_i = 0$ $X_i > R_i$ $P(X_i > R_i) = 1 - F(R_i;\theta)$

$$L(\theta \mid \underset{\sim}{Y}) = \prod_{i=1}^{n} \left[f_X(Y_i;\theta)\right]^{\delta_i} \left[1 - F(Y_i;\theta)\right]^{1-\delta_i}$$

$$f_X(x,\theta) = \frac{1}{\theta}\exp\left(\frac{x}{\theta}\right). \quad \theta > 0 \quad F_X(x;\theta) = 1 - \exp\left(\frac{x}{\theta}\right)$$

$$L(\theta \mid \underset{\sim}{Y}) = \prod_{i=1}^{n} \left[\frac{1}{\theta}\exp\left(\frac{Y_i}{\theta}\right)\right]^{\delta_i} \left[\exp\left(\frac{Y_i}{\theta}\right)\right]^{1-\delta_i}$$

$$= \left(\frac{1}{\theta}\right)^{\Sigma\delta_i} \prod_{i=1}^{n} \exp\left(\frac{\delta_i Y_i}{\theta} + \frac{(1-\delta_i)Y_i}{\theta}\right)$$

$$= \left(\frac{1}{\theta}\right)^{\Sigma\delta_i} \prod_{i=1}^{n} \exp\left(\frac{Y_i}{\theta}\right)$$

$$l(\underset{\sim}{\theta}|\underset{\sim}{Y}) = \log L(\underset{\sim}{\theta}|\underset{\sim}{Y}) = -\Sigma\delta_i \log\theta + \frac{\Sigma Y_i}{\theta}$$

$$\hat{\theta} = \frac{\Sigma Y_i}{\Sigma\delta_i}$$

random

right censoring $\quad X_1, \cdots, X_n \overset{iid}{\sim} f_X(x;\theta)$ independent

random censoring times $R_1, \cdots, R_n \overset{iid}{\sim} G(\cdot)$

observed data $\begin{cases} Y_i = \min(X_i, R_i) \\ \delta_i = I(X_i \le R_i) \end{cases}$ $\quad \delta_i = 1 \Rightarrow Y_i = X_i \Rightarrow X_i \le R_i$

random random

$$\frac{P(Y_i \in (y-h, y+h), \delta_i = 1)}{2h} = \frac{P(X_i \in (y-h, y+h), X_i \le R_i)}{2h}$$

$$= \frac{1}{2h} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} I(y-h < x < y+h, x < r) \times \underbrace{f(x;\theta)\, g(r)}_{X \text{ and } R \text{ are independent}} dx\,dr$$

$$f(x,R) = f(x;\theta) \times g(r)$$

$$= \frac{1}{2h} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} I(y-h<x<y+h) \times I(x<r)\, f(x;\theta)\, g(r)\, dx\,dr$$

$$= \frac{1}{2h} \int_{-\infty}^{+\infty} I(y-h<x<y+h) f(x;\theta) \left[\int I(x<r)\, \underline{g(r)}\, dr\right] dx$$

$$= \frac{1}{2h} \int_{y-h}^{y+h} f(x;\theta) \times \left[\int_{x}^{+\infty} g(r)\, dr\right] dx$$

$$= \frac{1}{2h} \int_{y-h}^{y+h} f(x;\theta)\, [1 - G(x)]\, dx$$

$$h \to 0 \quad \to \quad (1 - G(y)) \cdot f(y;\theta)$$

$\delta_i = 0 \Rightarrow \begin{cases} Y_i = R_i \\ X_i > R_i \end{cases}$

Similarly $\dfrac{P(Y_i \in (y-h, y+h),\ \delta_i = 0)}{2h} \longrightarrow [1-F(y;\theta)]\, g(y)$

$$L(\underset{\sim}{\theta} \mid \underset{\sim}{Y}, \underset{\sim}{\delta}) = \prod_{i=1}^{n} \left[\{1-G(y_i)\} \times f(y_i;\theta)\right]^{\delta_i} \times \left[[1-F(y_i;\theta)]\, g(y_i)\right]^{1-\delta_i}$$

$$= \prod_{i=1}^{n} \left[f(y_i;\theta)\right]^{\delta_i} \{1-F(y_i;\theta)\}^{1-\delta_i}$$

$$\times \prod_{i=1}^{n} \{1-G(y_i)\}^{\delta_i} \{g(y_i)\}^{1-\delta_i}$$

(Y_i, X_i) $\quad Y_i \sim X_i$ $\quad Y_i = X_i\beta + \varepsilon_i$

$$L(\theta \mid Y_i, \underset{\sim}{X}) = \prod_{i=1}^{n} f_{y,x}(y_i, x_i)$$

$$= \prod_{i=1}^{n} f_{y|x}(y_i|x_i)\, f_x(x_i)$$

$$= \boxed{\prod_{i=1}^{n} f_{y|x}(y_i|x_i;\theta)} \times f_x(x_i)$$

$\beta, \sigma^2, \ldots$

(Y_i, X_i) $\quad i = 1, \cdots, N$ $\qquad Y_i = \beta_0 + X_i\beta_1 + \varepsilon_i$

1st group $(Y_{11}, X_{11}) \cdots (Y_{1C}, X_{1C}) \longrightarrow \bar{Y}_1 = \frac{1}{C}\sum_{i=1}^{C} Y_{1i} \quad \bar{X}_1 = \frac{1}{C}\sum_{i=1}^{C} X_{1i}$

$\vdots$

J group $(Y_{J1}, X_{J1}) \cdots (\cancel{X}_{JC}, X_{JC}) \longrightarrow \bar{Y}_J \quad \cdots \quad \bar{X}_J$

$$\bar{Y}_j = \beta_0 + \bar{X}_j \beta_1 + \underbrace{\frac{1}{C}\sum_{i=1}^{C} \varepsilon_i}_{\bar{\varepsilon}_j}$$

$$\log Y_{ij} = \beta_0 + X_{ij}\beta_1 + \varepsilon_{ij}$$

Like $\rightarrow$ $(\bar{Y}_j, \bar{X}_j)$? estimate..

$$\log \bar{Y}_j \neq \beta_0 + \bar{X}_j \beta_1 + \bar{\varepsilon}_j$$

$f(y_{ij}, x_{ij})$

Example. The annual maximum sea levels in Venice 1931-1981

$$Y_i = x_i\beta + e_i \qquad e_i \sim N(0, \sigma^2)$$

$$Y_i = x_i\beta + e_i \qquad e_i \sim f_e(t, \sigma)$$

$$f_e(t) = \exp(-t)\exp\{-\exp(-t)\} \qquad t > 0$$

extreme value density

$$\frac{Y_i - x_i\beta}{\sigma} \sim f_e(t)$$