STA 380

Week 1

1. Let X have the probability density function:

$$f_X(x) = \begin{cases} 3e^{-3x} & x > 0, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Compute $\mathbb{E}[X]$.
(b) Compute $\mathbb{E}[e^{2x}]$.
(c) Compute $\mathbb{V}[X]$.

Exponential Distribution

(1) $f(y) = \frac{1}{\beta} e^{-y/\beta}$; $\beta > 0$, $0 \le y < \infty$

Mean $= \beta$

Variance $= \beta^2$

(2) $f(y) = \lambda \cdot e^{-\lambda y}$ ← rate

$\lambda > 0$

$0 \le y < \infty$

Mean $= \frac{1}{\lambda}$

Variance $= \frac{1}{\lambda^2}$

$f_X(x) = 3e^{-3x}$, $x > 0$

$\frac{1}{\beta} = 3 \Rightarrow \beta = \frac{1}{3}$

(a) $E[X] = \beta = \frac{1}{3}$

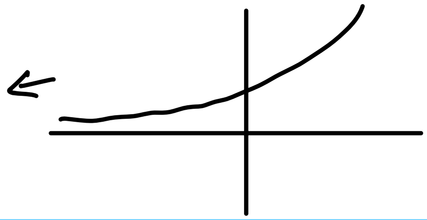
(c) $Var[X] = \beta^2 = (\frac{1}{3})^2 = \frac{1}{9}$

(b) $E[e^{2x}] = \int e^{2x} \cdot f(x)\, dx$

$$= \int_0^\infty e^{2x} \cdot 3e^{-3x}\, dx$$

$$= \int_0^\infty 3 \cdot e^{-x} dx$$

$$= 3\int_0^\infty e^{-x} dx = -3e^{-x}\Big|_0^\infty = -3(e^{-\infty} - e^0)$$



$= -3(0-1) = -3 \cdot -1$

$= 3$

2. Suppose we have a sample of i.i.d. random variables $X_1, X_2, \ldots, X_n$ with the following probability density function:

$$f_X(x) = \frac{2}{\theta} e^{-\frac{2}{\theta}x}, \quad x > 0, \ \theta > 0.$$

Furthermore, suppose my estimator for θ is $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

(a) Compute $\mathbb{E}[\overline{X}]$, the **expected value** of the estimator.

(b) Compute $\mathbb{V}[\overline{X}]$, the **variance** of the estimator.

(c) Compute $\mathbb{B}[\overline{X}]$, the **bias** of the estimator. Is this estimator unbiased?

(d) Compute $MSE[\overline{X}]$, the **mean squared error** of the estimator.

$$f_X(x) = \lambda e^{-\lambda x}$$

Mean $= \frac{1}{\lambda}$

Variance $= \frac{1}{\lambda^2}$

$$\lambda = \frac{2}{\theta}$$

(a) $E[\overline{X}] = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n} E(X_1 + X_2 \cdots + X_n)$

$$= \frac{1}{n}\left(E(X_1) + E(X_2) + \cdots E(X_n)\right)$$

$X_i \overset{iid}{\sim} Exp\left(\frac{2}{\theta}\right)$

$$= \frac{1}{n}\sum_{i=1}^{n} E(X_i)$$

$$= \frac{1}{\cancel{n}} \cdot \cancel{n} \cdot E(X_i) = \frac{1}{\frac{2}{\theta}} = \frac{\theta}{2}$$

(b)

$$Var(\overline{X}) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)$$

$$\overset{iid}{=} \frac{1}{n^2} \cdot \sum_{i=1}^{n} Var(X_i)$$

$$= \frac{1}{n^2} \cdot n \cdot \frac{1}{\left(\frac{2}{\theta}\right)^2} = \frac{1}{n} \cdot \frac{\theta^2}{4} = \frac{\theta^2}{4n}$$

Review:

$Var(aX)$
$= a^2 \cdot Var(X)$

(c)

Review:

$E[\overline{X}] - \theta = 0$

$$B[\overline{X}] = E[\overline{X}] - \theta = \frac{\theta}{2} - \theta = -\frac{\theta}{2} \neq 0$$

This is not an unbiased estimator.

(d) $MSE[\bar{X}] = Var[\bar{X}] + B[\bar{X}]^2 = \frac{\theta^2}{4n} + (-\frac{\theta}{2})^2$

$$= \frac{\theta^2}{4n} + \frac{\theta^2 n}{4n}$$

$$= \frac{\theta^2(1+n)}{4n}$$

can be proved

$$MSE[\bar{X}] = E[(\bar{X}-\theta)^2]$$

$$= E[\bar{X}^2 - 2\bar{X}\theta + \theta^2]$$

$$= E[\bar{X}^2] - 2\theta E[\bar{X}] + \theta^2$$

$$= Var(\bar{X}) + (E(\bar{X}))^2 - 2\theta E[\bar{X}] + \theta^2$$

$$= \frac{\theta^2}{4n} + (\frac{\theta}{2})^2 - \underbrace{2\theta \cdot \frac{\theta}{2}}_{\theta^2} + \theta^2 = \frac{\theta^2}{4n} + \frac{\theta^2 n}{4n}$$

$$= \frac{\theta^2(1+n)}{4n}$$

Review:

$Var(Y) = E(Y^2) - (E(Y))^2$

$\Rightarrow Var(Y) + (E(Y))^2 = E(Y^2)$

3. Suppose $n = 100$ people attended a screening of the highly anticipated Minecraft movie. For various reasons, some viewers chose to leave before it ended. The average and variance of the time spent watching the movie were 40 minutes and 5 minutes, respectively. Construct a 90% confidence interval to estimate μ, the true mean time at which viewers decide to leave the movie. Use the pivotal quantity:

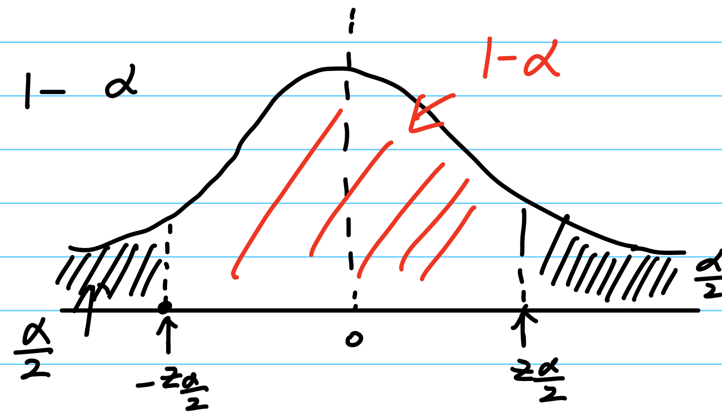
$$Z = \frac{\overline{x} - \mu}{\sigma_{\overline{x}}}.$$

$n = 100$

mean = 40 mins

variance = 5 mins

$$P(-z_{\frac{\alpha}{2}} \le Z \le z_{\frac{\alpha}{2}}) = 1 - \alpha$$



$$= P\left(-z_{\frac{\alpha}{2}} \le \frac{\overline{x} - \mu}{\sigma_{\overline{x}}} \le z_{\frac{\alpha}{2}}\right)$$

$$= P(-z_{\frac{\alpha}{2}} \cdot \sigma_{\overline{x}} \le \overline{x} - \mu \le z_{\frac{\alpha}{2}} \cdot \sigma_{\overline{x}})$$

$$= P(\underbrace{-\overline{x} - z_{\frac{\alpha}{2}} \cdot \sigma_{\overline{x}}}_{-(\overline{x} + z_{\frac{\alpha}{2}} \cdot \sigma_{\overline{x}})} \le -\mu \le \underbrace{-\overline{x} + z_{\frac{\alpha}{2}} \cdot \sigma_{\overline{x}}}_{= -(\overline{x} - z_{\frac{\alpha}{2}} \cdot \sigma_{\overline{x}})})$$

$$= P(\overline{x} + z_{\frac{\alpha}{2}} \cdot \sigma_{\overline{x}} \ge \mu \ge \overline{x} - z_{\frac{\alpha}{2}} \cdot \sigma_{\overline{x}})$$

$$= P(\overline{x} - z_{\frac{\alpha}{2}} \cdot \sigma_{\overline{x}} \le \mu \le \overline{x} + z_{\frac{\alpha}{2}} \cdot \sigma_{\overline{x}}) = 1 - \alpha = 90\% = 0.9$$

$$\Rightarrow \alpha = 0.1$$

90% CI $\quad \overline{x} \pm z_{\frac{\alpha}{2}} \cdot \sigma_{\overline{x}}$

$$\mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \overset{iid}{=} \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}(X_i) = \frac{1}{n^2} \cdot n \cdot \sigma^2 = \frac{\sigma^2}{n}$$

$$\sigma_{\overline{x}} = \sqrt{\mathrm{Var}(\overline{x})} = \sqrt{\frac{\mathrm{Var}(X)}{n}}$$

$$= \sqrt{\frac{5}{100}}$$

4. Let $Y_1, Y_2, ..., Y_n$ denote a random sample from a distribution with the following probability density function with parameters $\alpha > 0$ and $\beta > 0$, where β is known.

$$f(y|\alpha,\beta) = \begin{cases} \alpha\beta^{\alpha}y^{-(\alpha+1)} & y \geq \beta \\ 0 & \text{otherwise} \end{cases}$$

Find the MLE for α.

Step 1: $L(y_1 \cdots y_n|\alpha) = \alpha \cdot \beta^{\alpha} \cdot y_1^{-(\alpha+1)} \cdot \alpha \cdot \beta^{\alpha} \cdot y_2^{-(\alpha+1)} \cdots\cdots \alpha \cdot \beta^{\alpha} \cdot y_n^{-(\alpha+1)}$

$$= \alpha^{n} \cdot \beta^{\alpha n} \cdot \prod_{i=1}^{n} y_i^{-(\alpha+1)}$$

Step 2: Log-likelihood

$$\ell(y_1 \cdots y_n|\alpha) = \ln\left(\alpha^{n} \cdot \beta^{\alpha n} \cdot \prod_{i=1}^{n} y_i^{-(\alpha+1)}\right)$$

$$= \ln(\alpha^{n}) + \ln(\beta^{\alpha n}) + \ln\left(\prod_{i=1}^{n} y_i^{-(\alpha+1)}\right)$$

Review:

① $\ln(x^{y}) = y\ln(x)$

② $\ln(x \cdot y) = \ln(x) + \ln(y)$

$$= n \cdot \ln(\alpha) + \alpha n \cdot \ln(\beta) + \sum_{i=1}^{n} \ln\left(y_i^{-(\alpha+1)}\right)$$

$$= n \cdot \ln(\alpha) + \alpha n \cdot \ln(\beta) - (\alpha+1)\sum_{i=1}^{n} \ln(y_i)$$

Step 3: $\ell'(y_1 \cdots y_n|\alpha) = \frac{n}{\alpha} + n \cdot \ln(\beta) - \sum_{i=1}^{n} \ln(y_i) \overset{\text{set}}{=} 0$

$$\Rightarrow \frac{n}{\alpha} = \sum_{i=1}^{n} \ln(y_i) - n \cdot \ln(\beta)$$

$$\Rightarrow \hat{\alpha} = \frac{n}{\sum_{i=1}^{n} \ln(y_i) - n\ln(\beta)}$$

$\frac{n}{\alpha} = n \cdot \alpha^{-1}$

Step 4: $\ell'' = n \cdot -1 \cdot \alpha^{-2} = \frac{-n}{\alpha^{2}} < 0$

concave down

$n > 0$

$\alpha^{2} > 0$

always maximum

5. Let $Y_1, \ldots, Y_n$ be a random sample with the following common probability mass function:

$$f(y) = \begin{cases} \theta(1-\theta)^{y-1} & y = 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases}$$

Here, the unknown parameter $0 < \theta < 1$.

(a) Find the MLE of θ.

(b) Find the MLE of $\mathbb{E}(Y_1)$.

(a) skip $\hat{\theta} = \frac{1}{\bar{y}}$

(b) Invariance property of MLE:

if $\hat{\theta}$ is the MLE for θ and any injective function $g(x)$,

the MLE of $g(\theta)$ is $g(\hat{\theta})$

Geometric Distribution.

$P(y) = p(1-p)^{y-1}$ $y = 1, 2, \ldots\ldots$

Mean $\frac{1}{p}$ Variance $\frac{1-p}{p^2}$

$E(Y_1) = \frac{1}{\theta}$ injective for $0 < \theta < 1$

$\hat{\theta} = \frac{1}{\bar{y}}$

$\frac{1}{\hat{\theta}} = \frac{1}{\frac{1}{\bar{y}}} = \bar{y}$ is the MLE of $E[Y_1]$