C.1 Parameter Estimation

Sources

This lesson comes from Hogg McKean & Craig (7th Ed.) section 4.1 and Wackerly sections 8.1-8.3.

Functions of Samples

The number of customers entering a store between 9 and 10 A.M. is recorded on 3 days to be 9, 15, and 11.

$X =$ the number of customers on a day $\sim F(x) = \mathrm{P}[X \leqslant x]$

What we can learn about F depends on what we assume.

1. No assumption $\longrightarrow$ nonparametric statistics (later)
2. Assume some shape, and maybe some parameter values.

Examples: $X \sim$ Poisson(λ), $N(\mu, \sigma^2)$, $N(\mu, 25)$

A random sample of size 3 (Independent): $\mathbb{X} = (X_1, X_2, X_3)$, $X_i \sim X$.

A realization of the sample: $(9, 15, 11)$, (x_1, x_2, x_3)

$T(\mathbb{X}) = \frac{X_1+X_2+X_3}{3}$ is a statistic

$T(\mathbb{X} = (9, 15, 11)) = \frac{9+15+11}{3} = \frac{35}{3} = t$ is a realization of T.

Estimators

The number of customers entering a store between 9 and 10 A.M. is recorded on 3 days to be 9, 15, and 11.

Suppose $\mathbb{X} = (X_1, \ldots, X_n)$ with $X_i \sim$ Poisson(λ)

$\hat{\lambda} = T$ is an estimator of λ

$\hat{\lambda}(\mathbb{X}) = T(\mathbb{X})$ is random. $\hat{\lambda}(\mathbb{X} = (x_1, \ldots, x_n))$ is a number.

$\hat{\lambda}$ is notationally tricky. Sometimes a function, sometimes a number.

Possible choices for estimators:

$\hat{\lambda}_1 = \bar{X}$, $\quad \hat{\lambda}_2 = X_1$, $\quad \hat{\lambda}_3 = 12$, $\quad \hat{\lambda}_4 = \bar{X} + 1/n$

For this data set,
$\hat{\lambda}_1 = 35/3$, $\quad \hat{\lambda}_2 = 9$, $\quad \hat{\lambda}_3 = 12$, $\quad \hat{\lambda}_4 = 35/3 + 1/3 = 12$.

Expected Value of an Estimator

$$\hat\lambda_1 = \bar X \qquad \hat\lambda_2 = X_1, \qquad \hat\lambda_3 = 12, \qquad \hat\lambda_4 = \bar X + 1/n$$

$$\mathsf{E}[\hat\lambda]?$$

$$\mathsf{E}[\hat\lambda] = \sum_{\mathbb{X}=(x_1,\ldots,x_n)} \hat\lambda(\mathbb{X})\mathrm{P}[\mathbb{X} = (x_1,\ldots,x_n)]$$

$$\begin{aligned}\mathsf{E}[\hat\lambda_3] &= \sum 12\,\mathrm{P}[\mathbb{X} = (x_1,\ldots,x_n)] \\ &= 12\sum \mathrm{P}[\mathbb{X} = (x_1,\ldots,x_n)] = 12\cdot 1\end{aligned}$$

$$\mathsf{E}[\hat\lambda_2] = \mathsf{E}[X_1] = \lambda$$

$$\mathsf{E}[\hat\lambda_1] = \mathsf{E}\left[\left(\sum X_i\right)/n\right] = (1/n)\sum E[X_i] = n\lambda/n = \lambda$$

$$\mathsf{E}[\hat\lambda_4] = \mathsf{E}\left[\left(\sum X_i\right)/n + 1/n\right] = \lambda + 1/n$$

Variance of an Estimator

$$\hat\lambda_1 = \bar X \qquad \hat\lambda_2 = X_1, \qquad \hat\lambda_3 = 12, \qquad \hat\lambda_4 = \bar X + 1/n$$

$$\mathrm{Var}[\hat\lambda] = \mathsf{E}[\hat\lambda^2] - \mathsf{E}[\hat\lambda]^2, \text{ but sometimes } \ldots$$

$$\mathrm{Var}[\hat\lambda_3] = \mathrm{Var}[12] = 0$$

$$\mathrm{Var}[\hat\lambda_2] = \mathrm{Var}[X_1] = \lambda$$

$$\mathrm{Var}[\hat\lambda_1] = \mathrm{Var}\left[(1/n)\sum X_i\right] = (1/n)^2\sum \mathrm{Var}(X_i) = n\lambda/n^2 = \lambda/n.$$

$$\mathrm{Var}(X_1 + X_2) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + 2(\mathrm{Cov}(X_1, X_2) = 0).$$

$$\mathrm{Var}[\hat\lambda_4] = \mathrm{Var}\left[(1/n)\sum X_i + 1/n\right] = \lambda/n$$

Exercise 1

Suppose that $X_1, \ldots, X_n$ is a random sample, and $X_i \sim N(\mu, \sigma^2)$.

With $\hat{\mu} = \bar{X} + 1/n$ as the estimator for μ, what is the variance of $\hat{\mu}$?

Exercise 1

Suppose that $X_1, \ldots, X_n$ is a random sample, and $X_i \sim N(\mu, \sigma^2)$.

With $\hat{\mu} = \bar{X} + 1/n$ as the estimator for μ, what is the variance of $\hat{\mu}$?

$$\begin{aligned}
Var(\hat{\mu}) &= Var\left(\frac{1}{n} \sum X_i + \frac{1}{n}\right) \\
&= \left(\frac{1}{n}\right)^2 \sum Var(X_i) + 0 \\
&= \left(\frac{1}{n}\right)^2 n \sigma^2 \\
&= \boxed{\sigma^2 / n}
\end{aligned}$$

Exercise 2

Suppose that $X_1, \ldots, X_n$ is a random sample, and $X_i \sim N(\mu, \sigma^2)$.

Using $\hat{\sigma}^2 = \frac{\sum(X_i - \bar{X})^2}{n}$ as the estimator for σ^2, compute $E[\hat{\sigma}^2]$.

Exercise 2

Suppose that $X_1, \ldots, X_n$ is a random sample, and $X_i \sim N(\mu, \sigma^2)$.

Using $\hat{\sigma}^2 = \frac{\sum(X_i - \bar{X})^2}{n}$ as the estimator for σ^2, compute $E[\hat{\sigma}^2]$.

$$E[\hat{\sigma}^2] = E\left[\frac{\sum (X_i - \bar{X})^2}{n}\right]$$

$$= \frac{1}{n} E\left[\sum\left(X_i^2 - 2X_i\bar{X} + \bar{X}^2\right)\right]$$

$$= \frac{1}{n} E\left[\left(\sum X_i^2\right) - 2n\bar{X}\frac{\sum X_i}{n} + n\bar{X}^2\right]$$

$$\frac{1}{n}\left(\left(\sum E[X_i^2]\right) - E[n\bar{X}^2]\right)$$

C.1 Parameter Estimation C.1.1 Estimator Basics 7 / 7

Exercise 2 - continued

$$\frac{1}{n}\left(\left(\sum E[X_i^2]\right) - E\left[n\bar{X}^2\right]\right) \qquad X_i \sim N(\mu, \sigma^2)$$

$$E[X_i^2] = \sigma^2 + \mu^2$$

$$\frac{1}{n}\left(n(\sigma^2+\mu^2) - n\left(\operatorname{Var}\bar{X} + (E[\bar{X}])^2\right)\right) \qquad E[\bar{X}^2] = \operatorname{Var}(\bar{X}) + (E[\bar{X}])^2$$

$$\sigma^2 + \cancel{\mu^2} - \sigma^2/n - \cancel{\mu^2}$$

$$E[\hat{\sigma}^2] = \frac{n\sigma^2}{n} - \frac{1\sigma^2}{n} = \boxed{\left(\frac{n-1}{n}\right)\sigma^2}$$