Estimates and Estimators

Data Science for Quality Management: Sampling Distributions, Error and Estimation

with **Wendy Martin**

Learning objectives:

Describe types of estimates

List the criteria for “good” estimators

Types of Estimates

Point Estimate

- A single number used to estimate an unknown parameter

Types of Estimates

Interval Estimate

- A range of values used to estimate a population parameter

Types of Estimates

Estimator

- A sample statistic used to estimate a population parameter. An estimate is a specific observed value of a statistic.

Criteria for “Good” Estimators

- Unbiased
- Efficient
- Consistent
- Sufficient

Criteria for “Good” Estimators

Unbiasedness

- The mean of the Random Sampling Distribution (RSD) of the estimator is equal to the parameter it estimates.

Criteria for “Good” Estimators

Efficiency

- Refers to the standard error of the statistic RSD. The most efficient estimator is the one with the smallest standard error.

Criteria for “Good” Estimators

Consistency

- Refers to the assumption that as n increases, the value of the statistic approaches the value of its associated population parameter.

Criteria for “Good” Estimators

Sufficiency

- Refers to using all possible information in the sample to estimate the corresponding parameter.

Point Estimates

Point Estimate		Population Parameter	
Sample Mean	$\bar{X}$	Population Mean	μ
Sample Variance	s^2	Population Variance	σ^2
Sample Proportion	p	Population Proportion	π
Sample Count	c	Population Count	λ
Sample Skewness	g_3	Population Skewness	γ_3
Sample Kurtosis	g_4	Population Kurtosis	γ_4

Point Estimates

$$\bar{X} \approx \mu$$
$$s \approx \sigma \; and \; s^2 \approx \sigma^2$$
$$p \approx \pi$$
$$c \approx \lambda$$

Estimating σ from Multiple Samples

Average Range

$$\hat{\sigma} = \frac{\bar{R}}{d_2}$$

Median Range

$$\hat{\sigma} = \frac{\tilde{R}}{\tilde{d}_2}$$

Average Standard Deviation

$$\hat{\sigma} = \frac{\bar{s}}{c_4}$$

Median Standard Deviation

$$\hat{\sigma} = \sqrt{\frac{(\tilde{s})^2}{\chi_{0.5,n-1df}}}$$

Estimating σ from Multiple Samples

Average Variance (equal sample size)

$$\hat{\sigma} = \sqrt{\overline{s^2}}$$

(unequal sample size)

$$\hat{\sigma} = \sqrt{\frac{\sum_{j=1}^{k}(n_j - 1)s^2}{\sum_{j=1}^{k}(n_j - 1)}}$$

Estimating σ from Multiple Samples

Average Moving Range of the Mean

$$\hat{\sigma} = \frac{\overline{MR}_{\bar{X}}\sqrt{n}}{d_2}$$

Median Moving Range of the Means

$$\hat{\sigma} = \frac{\widetilde{MR}_{\bar{X}}\sqrt{n}}{\tilde{d}_2}$$

Standard Deviation of the Means $\hat{\sigma} = s_{\bar{X}}\sqrt{n}$

Point Estimates in RStudio

Point Estimate	In RStudio
$\bar{X}$	`mean()`
s	`sd()`
p	`mean() # average proportion`
c	`mean() # average count per unit`

Sources

- Luftig, J. An Introduction to Statistical Process Control & Capability. Luftig & Associates, Inc. Farmington Hills, MI, 1982