MATH 324: Statistics

Julian Lore

Last updated: January 23, 2018

Rough notes from Wackerly's Mathematical Statistics with Applications (7^{th} edition).

Contents

Information from Previous Chapters

Some formulas from probability:

$$Var(X) = E[(X-\mu)^2] = E[X^2] - (E[X])^2$$

Distributions:

Distribution	Probability Function	Mean	Variance	Moment-Generating Function
Binomial	$p(y) = \binom{n}{y} p^y (1-p)^{n-y}$	np	$np(1-p)$	$[pe^t + (1-p)]^n$

6.7 Order Statistics

Y_i, with $i = 1, \ldots, n$ independent continuous random variables. Denote ordered random variables by: $Y_{(1)} \leq Y_{(2)} \leq \ldots \leq Y_{(n)}$. We can then define:

$$Y_{(1)} = \min(Y_1, Y_2, \ldots, Y_n)$$
$$Y_{(n)} = \max(Y_1, Y_2, \ldots, Y_n)$$

In order to find the pdf of $Y_{(1)}$ and $Y_{(n)}$, we use the method of distribution functions.

$$\begin{aligned}F_{Y_{(n)}}(y) &= P(Y_{(n)} \leq y) = P(Y_1 \leq y, Y_2 \leq y, \dots, Y_n \leq y)\\ &\overset{\text{ind}}{=} P(Y_1 \leq y)P(Y_2 \leq y)\dots p(Y_n \leq y) = [F(y)]^n\end{aligned}$$

Derive for the density:

$$g_{(n)}(y) = n[F(y)]^{n-1}f(y)$$

$$\begin{aligned}F_{Y_{(1)}}(y) &= P(Y_{(1)} \leq 1) = 1 - P(Y_{(1)} > y) = 1 - 1 - P(Y_1 > y, Y_2 > y, \dots Y_n > y)\\ &\overset{\text{ind}}{=} 1 - [P(Y_1 > y)P(Y_2 > y)\dots P(Y_n) > y] = 1 - [1 - F(y)]^n\end{aligned}$$

Derive for the density:

$$g_{(1)}(y) = n[1 - F(y)]^{n-1}f(y)$$

Other Useful Information

$$E(\overline{Y}) = E(Y)$$

8 Estimation

8.1 Introduction

Point of statistics is to use sample information to infer data about the population. Populations are characterized by numbers (*parameters*) and we often want to estimate the value of parameter(s). Parameters include the proportion p, population mean μ, variance σ^2 and standard deviation σ.

Definition 1. The parameter of interest in an experiment is called the *target parameter*.

Definition 2. A *point estimate* is a type of estimate where we use a single value/point to estimate a parameter. If we estimate a parameter by saying that it might fall between two numbers, this is an *interval estimate.* We can use information from the sample to calculate these estimates, which are done using an estimator.

Definition 3. An *estimator* is a rule, often expressed as a formula, that tells how to calculate the value of an estimate based on the measurements contained in a sample.

Definition 4. *Sample mean:*

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

This is an example point estimator of μ.

There can be different estimators for the same population parameter. Some estimators are considered good and others are bad.

8.2 The Bias and Mean Square Error of Point Estimators

We cannot measure how good a point estimation procedure is with a single estimate, we need to observe the procedure many times. We create a frequency distribution to measure the goodness of a point estimator.

Point Estimators For a population parameter θ, the estimator of θ is called $\hat{\theta}$.

Definition 5. Ideally, we'd want $E(\hat{\theta}) = \theta$. Point estimators that satisfy this are called *unbiased.* Otherwise, they are called *biased*, where the *bias* is given by $B(\hat{\theta}) = E(\hat{\theta}) - \theta$

In addition, we'd also like the estimator $V(\hat{\theta})$ to be as small as possible, since a smaller variance guarantees a higher fraction of estimators to be "close" to θ. If two estimators are unbiased and everything else is equal other than variance, we prefer the one with smaller variance.

Definition 6. Another way to characterize goodness of a point estimator is via its *mean square error*,

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

Which is the average of the square of the distance between the estimator and its target parameter. It can be shown that:

$$MSE(\hat{\theta}) = V(\hat{\theta}) + [B(\hat{\theta})]^2$$