7 Limit Theorems

Motivation

For some new statistics, we may want to derive features of the distribution of the statistic.

When we can't do this analytically, we need to use statistical computing methods to *approximate* them.

We will return to some basic theory to motivate and evaluate the computational methods to follow.

7.1 Laws of Large Numbers

Limit theorems describe the behavior of sequences of random variables as the sample size increases ($n \to \infty$).

Often we describe these limits in terms of how close the sequence is to the truth.

We can evaluate this distance in several ways.

Some modes of convergence –

Laws of large numbers –

7.2 Central Limit Theorem

Theorem 7.1 (Central Limit Theorem (CLT)) Let $X_1, \ldots, X_n$ be a random sample from a distribution with mean μ and finite variance $\sigma^2 > 0$, then the limiting distribution of $Z_n = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}}$ is $N(0, 1)$.

Interpretation:

Note that the CLT doesn't require the population distribution to be Normal.

8 Estimates and Estimators

Let $X_1, \ldots, X_n$ be a random sample from a population.

Let $T_n = T(X_1, \ldots, X_n)$ be a function of the sample.

Statistics estimate *parameters*.

Example 8.1

Definition 8.1 An *estimator* is a rule for calculating an estimate of a given quantity.
Definition 8.2 An *estimate* is the result of applying an estimator to observed data samples
in order to estimate a given quantity.

We need to be careful not to confuse the above ideas:

$\overline{X}_n$

$\overline{x}_n$

μ

We can make any number of estimators to estimate a given quantity. How do we know the
"best" one?

9 Evaluating Estimators

There are many ways we can describe how good or bad (evaluate) an estimator is.

9.1 Bias

Definition 9.1 Let $X_1, \ldots, X_n$ be a random sample from a population, θ a parameter of interest, and $\hat{\theta}_n = T(X_1, \ldots, X_n)$ an estimator. Then the *bias* of $\hat{\theta}_n$ is defined as

$$bias(\hat{\theta}_n) = E[\hat{\theta}_n] - \theta.$$

Definition 9.2 An *unbiased estimator* is defined to be an estimator $\hat{\theta}_n = T(X_1, \ldots, X_n)$ where

Example 9.1

Example 9.2

Example 9.3

9.2 Mean Squared Error (MSE)

Definition 9.3 The *mean squared error (MSE)* of an estimator $\hat{\theta}_n$ for parameter θ is defined as

$$MSE(\hat{\theta}_n) = E\left[(\theta - \hat{\theta}_n)^2\right]$$
$$= Var(\hat{\theta}_n) + \left(bias(\hat{\theta}_n)\right)^2.$$

Generally, we want estimators with

Sometimes an unbiased estimator $\hat{\theta}_n$ can have a larger variance than a biased estimator $\tilde{\theta}_n$.

Example 9.4 Let's compare two estimators of σ^2.

$$s^2 = \frac{1}{n-1}\sum(X_i - \overline{X}_n)^2 \qquad \hat{\sigma}^2 = \frac{1}{n}\sum(X_i - \overline{X}_n)^2$$

9.3 Standard Error

Definition 9.4 The *standard error* of an estimator $\hat{\theta}_n$ of θ is defined as

$$se(\hat{\theta}_n) = \sqrt{Var(\hat{\theta}_n)}.$$

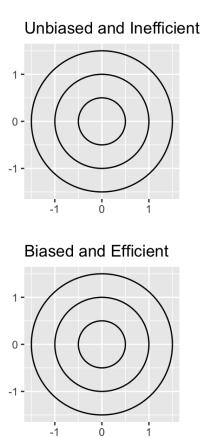We seek estimators with small $se(\hat{\theta}_n)$.

Example 9.5

10 Comparing Estimators

We typically compare statistical estimators based on the following basic properties:
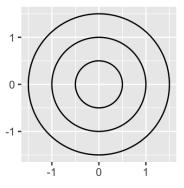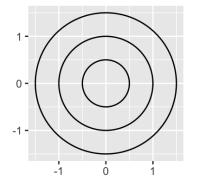
1.

2.

3.

4.

Unbiased and Inefficient

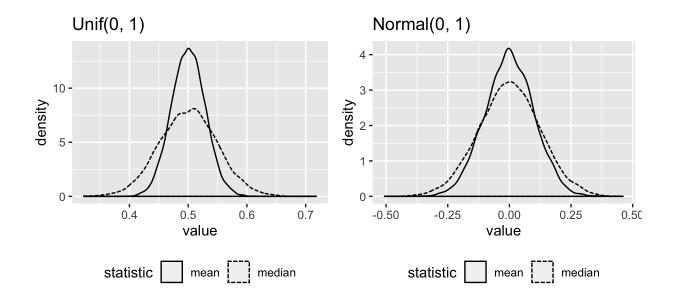### Biased and Inefficient

Biased and Efficient

Unbiased and Efficient

Example 10.1 Let us consider the efficiency of estimates of the center of a distribution. A **measure of central tendency** estimates the central or typical value for a probability distribution.

Mean and median are two measures of central tendency. They are both **unbiased**, which is more efficient?

```r
set.seed(400)

times <- 10000 # number of times to make a sample
n <- 100 # size of the sample
uniform_results <- data.frame(mean = numeric(times), median =
  numeric(times))
normal_results <- data.frame(mean = numeric(times), median =
  numeric(times))

for(i in 1:times) {
  x <- runif(n)
  y <- rnorm(n)
  uniform_results[i, "mean"] <- mean(x)
  uniform_results[i, "median"] <- median(x)
  normal_results[i, "mean"] <- mean(y)
  normal_results[i, "median"] <- median(y)
}

uniform_results %>%
  gather(statistic, value, everything()) %>%
  ggplot() +
  geom_density(aes(value, lty = statistic)) +
  ggtitle("Unif(0, 1)") +
  theme(legend.position = "bottom")

normal_results %>%
  gather(statistic, value, everything()) %>%
  ggplot() +
  geom_density(aes(value, lty = statistic)) +
  ggtitle("Normal(0, 1)") +
  theme(legend.position = "bottom")
```

Unif(0, 1)

Normal(0, 1)

Next Up In Ch. 5, we'll look at a method that produces *unbiased* estimators of $E(g(X))$!