MA22004

Seminar 2

Dr Eric Hall

10/15/2020

Announcements

Reminders

- It is week 2! You should have read §2 of the notes on **Perusall**.
- Feedback for Class Test (Practice) is posted.

Upcoming

- Lab 1 due **Sunday 18/10** at **13:00** (new deadline) upload to **Gradescope**.
- Worksheet #2 (Estimating Means) on **Blackboard** before next workshop.
- Investigation #1 on **Perusall** before next workshop.
- Reading assignment §3 on **Perusall** before next seminar.

But what are statistics?

A **statistic** is a quantity that can be calculated from sample data.

Prior to obtaining data, a statistic is an unknown quantity and is therefore a rv.

We refer to the probability distribution for a statistic as a sampling distribution to emphasize how the distribution will vary across all possible sample data.

Basics of statistical inference

Today we will look at three tools for drawing conclusions about the characteristics of a population from data:

- point estimation
- confidence intervals
- hypothesis testing

> We will move fast over topics discussed in Stat 1.

Point estimation

Consider iid $X_1, X_2, \ldots, X_m \sim F(\theta)$.

A **point estimator** $\widehat{\theta}_m$ of θ is obtained by selecting a suitable statistic g,

$$\widehat{\theta}_m = g(X_1, \ldots, X_m)\,.$$

Closing the deal: point estimate

A **point estimate** of a parameter θ is a single number that we regard as a sensible value for θ.

A point estimate $\widehat{\theta}_m$ is computed from an estimator using sample data.

> The symbol $\widehat{\theta}_m$ is typically used to denote *both* the estimator and the point estimate resulting from a given sample.

How uncertain is our point estimate?

The standard error is one measure of the precision of an estimate.

The **standard error** of an estimator $\widehat{\theta}$ is the standard deviation

$$\sigma_{\widehat{\theta}} = \sqrt{\mathrm{Var}(\widehat{\theta})}\,.$$

Often, the standard error depends on unknown parameters and must also be estimated.

The **estimated standard error** is denoted by $\widehat{\sigma}_{\widehat{\theta}}$ or simply $s_{\widehat{\theta}}$.

Confidence intervals

An interval estimate reports an entire range of plausible values for the parameter of interest.

A confidence interval is an interval estimate that makes a probability statement about the degree of reliability, or the confidence level, of the interval.

$100(1-\alpha)\%$ CI definition

A $100(1-\alpha)\%$ **confidence interval** for a parameter θ is a *random* interval $C_m = (L_m, U_m)$, where $L_m = \ell(X_1, \ldots, X_m)$ and $U_m = u(X_1, \ldots, X_m)$ are functions of the data, such that

$$P_\theta(L_m < \theta < U_m) = 1 - \alpha\,,$$

for all $\theta \in \Theta$ (the parameter space).

What a CI is not…

- A 95% confidence level does not mean there is a 95% probability that the population parameter lies within a given interval.
- A 95% confidence level does not mean that 95% of the sample data lie within the confidence interval.
- A particular confidence level of 95% calculated from an experiment does not mean that there is a 95% probability of a sample parameter from a repeat of the experiment falling within this interval.

Hypothesis testing

Methods for determining which of two contradictory claims, or **hypotheses**, about a parameter is correct.

The **null hypothesis**, denoted by H_0, is a claim that we initially assume to be true by default.

The **alternative hypothesis**, denoted by H_a, is an assertion that is contradictory to H_0.

Definition of a hypothesis test

A **hypothesis test** asks if the available data provides sufficient evidence to reject H_0.

If the observations disagree with H_0, then we reject the null hypothesis. On the other hand, if the sample evidence does not strongly contradict H_0, then we continue to believe H_0.

> The two possible conclustions of a hypothesis test are: *reject* H_0 or *fail to reject* H_0.

Know when to fold 'em

A **test statistic** T is a function of the sample data (like an estimator). The decision to reject or fail to reject H_0 will involve computing the test statistic.

The **rejection region** R is the collection of values of the test statistic for which H_0 is to be rejected in favor of the alternative, e.g.,

$$R = \{x : T(x) > c\} \ ,$$

where c is referred to as a **critical value**. If $X \in R$ then we reject H_0. The alternative is that $X \notin R$ and in this case we fail to reject the null.

Balancing act

The basis for choosing a rejection region involves balancing Type I and II errors. A conclusion is reached in a hypothesis test by selecting a **significance level** α for the test linked to the maximal type I error rate.

Knitting everything together

A P-**value** is the probability, calculated assuming H_0 is true, of obtaining a value of the test statistic at least as contradictory to H_0 as the value calculated from the sample data.

Smaller P-values indicate stronger evidence against H_0 in favor of H_a. If $P \leq \alpha$ then we reject H_0 at significance level α. If $P \geq \alpha$ we fail to reject H_0 at significance level α.

Anatomy of a hypothesis test

To summarize, the elements of a statistical test are:

1. Null hypothesis (H_0)
2. Alternative hypothesis (H_a)
3. Test statistic
4. Rejection region
5. Significance level (α)

Summary

Today we discussed:

- point estimation
- confidence intervals
- hypothesis tests