Hypothesis Testing for a Mean

Math 130: Introductory Statistics

July 6, 2009

1

Goals of this section

- Learn about what it means for statistical results to be statistically significant.
- Learn about how to use statistical results to test hypotheses.

2 Hypotheses

2.1 Constructing hypotheses

Hypothesis Test

- A **hypothesis** is a claim or statement about a property of a population.
  - Example: The population mean for systolic blood pressure is 120.
- A **hypothesis test** (or **test of significance** is a standard procedure for testing a claim about a property of a population.
- Recall the example from last week about birth weights with mothers who use drugs.
  - Hypothesis: Using drugs during pregnancy leads to an average birth weight of 7 pounds (the same as with mothers who do not use drugs).

Null and Alternative Hypotheses

- The **null hypothesis** is a statement that the value of a population parameter (such as the population mean) *is equal to* some claimed value.
  - H_0: $\mu = 7$.
- The **alternative hypothesis** is an alternative to the null hypothesis; a statement that says a parameter differs from the value given in the null hypothesis.

- H_a: $\mu < 7$.
- H_a: $\mu > 7$.
- H_a: $\mu \neq 7$.

- In hypothesis testing, assume the null hypothesis is true until there is strong statistical evidence to suggest the alternative hypothesis.

- Similar to an "innocent until proven guilty" policy.

3 Hypothesis test when σ is known

3.1 Computing significance

Test statistics

- A **test statistic** is a value computed with sample data that can be used to establish how far away a statistic (such as the sample mean) is from a population parameter given in the null hypothesis.

  - Z-statistic:

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

- Where μ is a value given in the null hypothesis.

Critical regions

- A **critical region** is a set of all values of the test statistic that will cause us to reject the null hypothesis.

- The critical region depends on the **significance level** given by α.

  - Common choices for α are 0.10, 0.05, 0.01.

- A **left tailed** test puts α is the left hand tail. Used to test alternative hypothesis with a *less than.*

- A **right tailed** test puts α is the left hand tail. Used to test alternative hypothesis with a *greater than.*

- A **two tailed** test puts α is the left hand tail. Used to test alternative hypothesis with a *not equal to.*

3.2 Type I and Type II errors

Type I and Type II errors

- **Type I error**: mistake of rejecting the null hypothesis when it is actually true. The value α is the probability of a type I error.
- **Type II error**: The mistake of failing to reject the null hypothesis when it is actually false.
- The probability of a type II error is usually given by β.
- The **power** of a test is the probability of *not* making a type II error, given by $1-\beta$.
- Power and type II probabilities are kind of complicated to compute.

3.3 Steps of a hypothesis test

Steps of a hypothesis test

- Construct null and alternative hypothesis.
- Given a value of α, find the critical region(s).
- Compute the test statistic.
- Compare the test statistic to the critical regions.
- Based on this comparison, reject or fail to reject the null hypothesis.
- Say (or write) in plain English the result you found.

Example:

- Lets construct a hypothesis test with $\alpha = 0.05$ to determine if using drugs leads to lower birth weights. Suppose $\bar{x} = 6$, $\sigma = 3.1$, $n = 50$.
- Hypotheses: H_0: $\mu = 7$. H_a: $\mu < 7$
- Get critical region. $z_{critical} = -1.645$ Critical region: $z < -1.645$.
- Compute test statistic. By central limit theorem can use z.

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{6-7}{3.1/\sqrt{50}} = -2.28.$$

- Since $z < z_{critical}$, reject the null hypothesis.
- There is sufficient statistical evidence to suggest that using drugs during pregnancy leads to lower birth weights.

3.4 P-value method

P-value method

- The **p-value** method is an alternative to finding critical regions for the z-statistic.
- Compute the test statistic:

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

- Right (left) tailed test: find the area to the right (left) of the test statistic (this is called the **p-value**).
- Compare the area to the value of α.
  - If p-value $< \alpha$, in rejection region.
  - If p-value $> \alpha$, not in the rejection region.
- Two tailed test: p-value = double the area in the tail beyond the test statistic.

4 T-distribution

4.1 Assumptions

When σ is unknown.

- Assume we have a sample size $n > 30$ *or..*
- Assume the population has a normal distribution.
- The population standard deviation, σ is not known. Our best estimate is the sample standard deviation, s.

4.2 T-distribution

More uncertainty

- Since we need to estimate σ with s, this adds more uncertainty.
- More uncertainty means we need to widen the confidence interval.
- Introduce a new critical statistic, the T-statistic (instead of using Z).
- The degree of uncertainty depends on the sample size.

T-distribution

- Recall the z-score was given by,

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

- The t-statistic only differs in that you use s instead of σ:

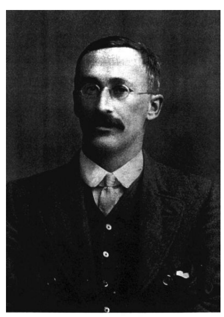
$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

- This has a **T-distribution** with $n - 1$ **degrees of freedom**.
- This depends on the sample size because the uncertainty about σ decreases with larger sample sizes.
- **Degrees of freedom**: the sample size minus the number of parameters you are estimating.

4.3 William Sealy Gosset

Ever wonder who came up with this?

- William Sealy Gosset (1876-1937) published these results in 1908.
- A quality control statistician employed by Guinness.
- Guinness prohibited all employees from publishing their research.



'Student' in 1908

5

5.1 Homework

Homework

- Hypothesis testing when σ is known: 7-12 odd.
- Hypothesis testing when σ is unknown: 13-18 odd.