Testing Hypotheses

Statistical Hypothesis Testing

- Statistical hypothesis testing is a procedure that allows us to evaluate hypotheses about population parameters based on sample statistics.

Assumptions of Statistical Hypothesis Testing

- The sample is a random sample.
- The level of measurement is interval-ratio.
- The population is normally distributed or that the sample size is larger than 50.

Assumptions of Statistical Hypothesis Testing

- The GSS looked at African Americans and their incomes. The sample was 100 respondents.
- Does that example meet all of the assumptions of statistical hypothesis testing?

The Research Hypothesis

- The research hypothesis is a statement reflecting the substantive hypothesis. It is always expressed in terms of population parameters.
- Research hypotheses are always expressed in terms of population parameters because we are interested in making statements about population parameters based on our sample statistics.

The Research Hypothesis

- The research hypothesis is symbolized by H_1.
- H_1: The average wages of blacks are lower than the average wages of all Americans, say $ 28,985.
- We use μ_Y to represent the black population mean, so we can express it $H_1: \mu_Y < \$28,985$

The Research Hypothesis

- The research hypothesis specifies that the population is one of the following

Not equal to some specified value: $\mu_Y \neq$ some specified value

Greater than some specified value: $\mu_Y >$ some specified value

Less than some specified value: $\mu_Y <$ some specified value

The Null Hypothesis

- It is possible that there is no difference between the mean wages of blacks and the mean wages of all Americans.
- The perceived difference could be due to the fact that the particular sample happened to contain blacks with lower earnings. We can only at best, estimate the likelihood that the research hypothesis is true or false.

The Null Hypothesis

- Statisticians use a hypothesis that is the opposite of the research hypothesis to assess how likely the research hypothesis is correct.
- The null hypothesis, symbolized as H_0, contradicts the research hypothesis and usually states that there is no difference between the population mean and the specified value.

The Null Hypothesis

- Rather than directly testing the research hypothesis we test the null hypothesis.
- We hope to reject the null hypothesis in order to provide support for the research hypothesis.
- The null hypothesis is a statement of no difference, which contradicts the research hypothesis and is always expressed in terms of population parameters.

Research and Null Hypotheses

- One tail test – specifies the hypothesized direction
- Research Hypothesis:
  - H_1: $\mu_2 > \mu_1$
- The Null Hypothesis:
  - H_0: $\mu_2 = \mu_1$

One-Tailed Test

- One-tailed test is a type of hypothesis test that involves a directional hypothesis. It specifies that the values of one group are either larger, (right-tailed test) or smaller, (left-tailed test) than some specified population value.

Research and Null Hypotheses

- Two Tail – direction is not specified (more common)
- Research Hypothesis:
  - H_1: $\mu_2 \neq \mu_1$
- Null Hypothesis:
  - H_0: $\mu_2 = \mu_1$

Two-Tailed Test

- Sometimes, we believe there is a difference between groups, but we cannot anticipate the direction of that difference.
- When we do not have an idea of the direction in the research hypothesis, we conduct a two-tailed test.

Two – Tailed Test

- A two-tailed test is the type of hypothesis test that involves a non-direction research hypothesis. We are equally interested in whether the values are less than or greater than one another. The sample outcome may be located at both the low and high ends of the sampling distribution.

Probability Values and Alpha

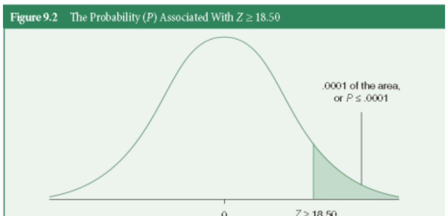
- We assume that our null hypothesis is true, and we want to determine whether our sample evidence casts doubt on that assumption. That actually, our research hypothesis is correct.
- What are the chances that we have randomly selected a sample of African American such that the average earnings are this much lower than the average for the general population.

Probability Values and Alpha

- We can determine the probability of our choosing such a sample because of what we know about the sampling distribution and its properties.
- We know based on the central limit theorem that if our sample size is larger than 50, the sampling distribution of the mean is approximately normal.

Probability Values

Figure 9.2 The Probability (P) Associated With $Z \geq 18.50$



Probability Values

- P value
  - The probability associated with the obtained value of Z
- Alpha (α)
  - The level of probability at which the null hypothesis is rejected
    - It is customary to set alpha at the .05, .01, or .001 level

Example

- We are assuming that the null hypothesis is true. We have a population mean μ_y = $28,985 and a standard deviation σ_y = $23,335.
- We have a sample size of 100, n = 100.
- So, we can assume that the sample mean is approximately normal.

Example

- Because this distribution of sample means is normal, we can use appendix B to determine the probability of drawing a sample mean of $24,100 or smaller from this population.
- We need to find the Z score before we go to appendix B.

Example

- Converting the sample mean to a Z score equivalent is called computing the test statistic.
- The Z value we obtain is called the Z statistic.
- The Z gives us the number of standard deviations that our sample is from the hypothesized value, assuming the null hypothesis is true.

Example

- First, we need to determine whether the Z score is consistent with our research hypothesis.
- Remember, we defined our research hypothesis as a left-tailed test. The negative value we obtained confirms that we will be evaluating the difference on the left tail.

Example

- We now take the Z score to appendix B to find the area of the negative Z of 2.09.
- Which column do you think we use once we go to appendix B?
- What is the probability of getting a result as extreme as the sample result if the null hypothesis is true?

Alpha

- Researchers usually define in advance what is a sufficiently improbable Z value by specifying a cutoff point below which P must fall to reject the null hypothesis.
- This cutoff is called alpha and is customarily set at the .05, .01. or .001 level.

Alpha

- Based on the P value, we can also make a statement regarding the significance of the results.
- If the P value is equal to or less than our alpha level, our obtained Z statistic is considered statistically significant. That is to say that it is very unlikely to have occurred by random chance or sampling error.

Z for a Two-Tailed Test

- To find P for a two-tailed test, look up the area in column C of appendix B that corresponds to your obtained Z and then multiply it by two to obtain the two –tailed probability.
- The two-tailed P value for Z = -2.09 is .0183 X 2 = .0366. Would we still be able to reject the null hypothesis?

The Five Steps to Hypothesis Testing

- Making assumptions
- Stating the research and null hypotheses and selecting alpha
- Selecting the sampling distribution and specifying the test statistic
- Computing the test statistic
- Making a decision and interpreting the results.

Type I and Type II Errors

- **Type I error** (false rejection error)—the probability (equal to α) associated with rejecting a true null hypothesis.
- **Type II error** (false acceptance error)—the probability associated with failing to reject a false null hypothesis.

Based on sample results, the decision made is to ...

		reject H_0	**do not reject H_0**
In the population H_0 is ...	**true**	**Type I error (α)**	correct decision
	false	correct decision	**Type II error**

The T Statistic

- In most situations, standard deviation will not be known as it was for the Z statistic.
- We then use the T statistic instead of the Z statistic to test the null hypothesis.

The T Statistic

- The formula for computing the T statistic is:

$$t = \frac{\bar{Y} - \mu_y}{\frac{S_y}{\sqrt{n}}}$$

T Distributions and Degrees of Freedom

- The t distribution is a family of curves, each determined by its degrees of freedom.
- The degrees of freedom (df) represent the number of scores that are free to vary in calculating each statistic.

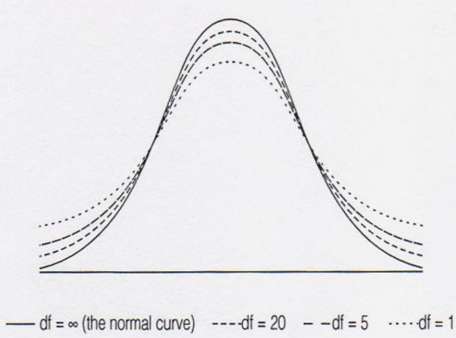
Calculating the Degrees of Freedom

- When calculating the degrees of freedom you must know the sample size and the whether there are any restrictions in calculating that statistic.
- Then the number of restrictions is subtracted from the sample size.

t Test

- **t statistic *(obtained)*** – The test statistic computed to test the null hypothesis about a population mean when the population standard deviation is unknown and is estimated using the sample standard deviation.
- **t distribution** – A family of curves, each determined by its degrees of freedom (df). It is used when the population standard deviation is unknown and the standard error is estimated from the sample standard deviation.
- **Degrees of freedom *(df)*** – The number of scores that are free to vary in calculating a statistic.

Figure 13.3 **The Normal Distribution and *t* Distributions for 1, 5, 20, and ∞ Degrees of Freedom**



T Distribution

- Appendix C summarizes the t distribution.
- The t table differs from the z table in several ways. The t table shows df, probabilities or alpha, significance levels are shown for both one-tailed tests and two-tailed tests.
- On the next slide find the t value of 2.021 with 40 df's and a two –tailed test.

Table 13.2 **Values of the *t* Distribution**

	Level of Significance for One-Tailed Test					
	.10	.05	.025	.01	.005	.0005
	Level of Significance for Two-Tailed Test					
df	**.20**	**.10**	**.05**	**.02**	**.01**	**.001**
1	3.078	6.314	12.706	31.821	63.657	636.619
2	1.886	2.920	4.303	6.965	9.925	31.598
3	1.638	2.353	3.182	4.541	5.841	12.941
4	1.533	2.132	2.776	3.747	4.604	8.610
5	1.476	2.015	2.571	3.365	4.032	6.859
10	1.372	1.812	2.228	2.764	3.169	4.587
15	1.341	1.753	2.131	2.602	2.947	4.073
20	1.325	1.725	2.086	2.528	2.845	3.850
25	1.316	1.708	2.060	2.485	2.787	3.725
30	1.310	1.697	2.042	2.457	2.750	3.646
40	1.303	1.684	2.021	2.423	2.704	3.551
60	1.296	1.671	2.000	2.390	2.660	3.460
120	1.289	1.658	1.980	2.358	2.617	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.291

Source: Abridged from R. A. Fisher and F. Yates, *Statistical Tables for Biological, Agricultural and Medical Research*, Table 111.

Example

- Let's try a two-tailed hypothesis out about a population mean.
- $\bar{Y}$ = \$26,078
- S_y = \$21,751
- μ_y = \$20,410
- σ_y = ?
- N = 150

Example

- H_1: $\mu_y \neq \$20,410$
- H_0: $\mu_y = \$20,410$
- We will set alpha as .05, meaning that we will reject the null hypothesis if the probability of our obtained statistic is less than or equal to .05.

Example

- First we need to calculate the df associated with our test.
- Since we only have one sample our restriction is one.
- What is the df?

Example

- Next we need to calculate the t statistic.
- The formula is:

$$t = \frac{\bar{Y} - \mu_y}{\frac{S_y}{\sqrt{n}}}$$

Example

- When we go to Appendix C we see that there is no row for 149 df's so we have to use the last row df = ∞ to locate our obtained t statistic.
- What is our significance level?
- Can we reject the null hypothesis?

t- tests for Difference Between Means

- The sampling distribution allows us to compare our sample results with all possible sample outcomes and estimate the likelihood of their occurrence.
- To get the t-test score we need the difference between the two means, the standard error, and the degrees of freedom.

Difference Between Means t-test

$$t = \frac{\bar{Y}_1 - \bar{Y}_2}{S_{\bar{y}_1} - S_{\bar{y}_2}}$$

Difference between the Means

The estimated standard error

t- tests for Difference Between Means if Population Variances appear equal

- The difference between the two means: $\bar{Y}_1 - \bar{Y}_2$
- The degrees of freedom:

$$df = (N_1 + N_2) - 2$$

The standard error:

$$S_{\bar{y}1} - S_{\bar{y}2} = \sqrt{\frac{(N_1 - 1)\, S^2_{y1} + (N_2 - 1) S^2_{y2}}{(N_1 + N_2) - 2}} \sqrt{\frac{N_1 + N_2}{N_1 N_2}}$$

t- tests for Difference Between Means if Population Variances appear unequal

- The difference between the two means: $\bar{Y}_1 - \bar{Y}_2$
- The degrees of freedom:

$$df = \frac{(S^2_{\bar{y}_1}/N_1 + S^2_{\bar{y}2}/N_2)^2}{(S^2_{\bar{y}_1}/N_1)^2/(N_1-1) + (S^2_{\bar{y}2}/N_2)^2/(N_2-1)}$$

The standard error:

$$S_{\bar{Y}_1-\bar{Y}_2} = \sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}$$

t-Test Example 1

Mean pay according to gender:

	N	Mean Pay	S.D.
Women	50	$10.29	.8766
Men	54	$10.06	.9051

What can we conclude about the difference in wages?