THE NORMAL DISTRIBUTION AND TESTS OF STATISTICAL SIGNIFICANCE

One sample Z test; test of proportions; and t-test

Question: What is the one sample Z-test? How does it differ from the Z-test for the comparison of group means?

WORKING WITH PROBABILITIES AND Z-SCORES (APPENDIX 3)

If X is **normally distributed** then its **probability density function** is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad \text{for} \quad -\infty < x < \infty$$

where μ is the mean and σ^2 is the variance of the distribution.

We write $\quad X \sim \mathrm{N}(\mu, \sigma^2)$.

Z-SCORES: THE STANDARDIZED NORMAL

All normal distributions can be converted to standard normal

Standard normal distributions have mean = 0 and s = 1

Why?

Single observation not a mean

Population mean

$$z = \frac{X_i - \mu}{\sigma}$$

Population standard deviation

A RANDOMLY SELECTED CHILD HAS AN IQ OF 124, HOW DOES THIS SCORE COMPARE TO THE GENERAL POPULATION?

IQ scores are normally distributed in the US with a mean of 100 and standard deviation of 15.

Transform the IQ score to a z-score and then compare this to the standard normal

$$z = \frac{X_i - \mu}{\sigma} = \frac{124 - 100}{15} = \frac{24}{15} = 1.60$$

DEVELOPING TESTS OF STATISTICAL SIGNIFICANCE BASED ON STANDARD NORMAL

The normal distribution can also be used as a sampling distribution

Let's say that you were interested in whether people who live in the West differ from people living in the US generally on average IQ scores.

<u>The population characteristics for all people in the US are known.</u> The mean score for the population is 100, and the standard deviation of the population mean is 15.

You conduct a study of 125 people living in California selected through an independent random sampling procedure from the population of all people who live in the state. You find that the mean IQ in your sample is 90.

What are you interested in knowing here? I.e., what hypothesis are you testing?

ONE SAMPLE Z-TEST

As before, we want to know if the observed IQ score is significantly different from what we know about the average IQ score in the population

Here the average in the population is 100, with known standard deviation

H0: The mean IQ of our sampled Californians is the same as the general population → $H_0\colon\ \mu = 100$

H1: The mean is different → $H_0\colon\ \mu \neq 100$

Because the population parameters of the American population are known we use a single-sample z-test

ASSUMPTIONS

Level of Measurement: Interval scale.

Population Distribution: Normal distribution.

Sampling Method: Independent random sampling.

Sampling Frame: The California population.

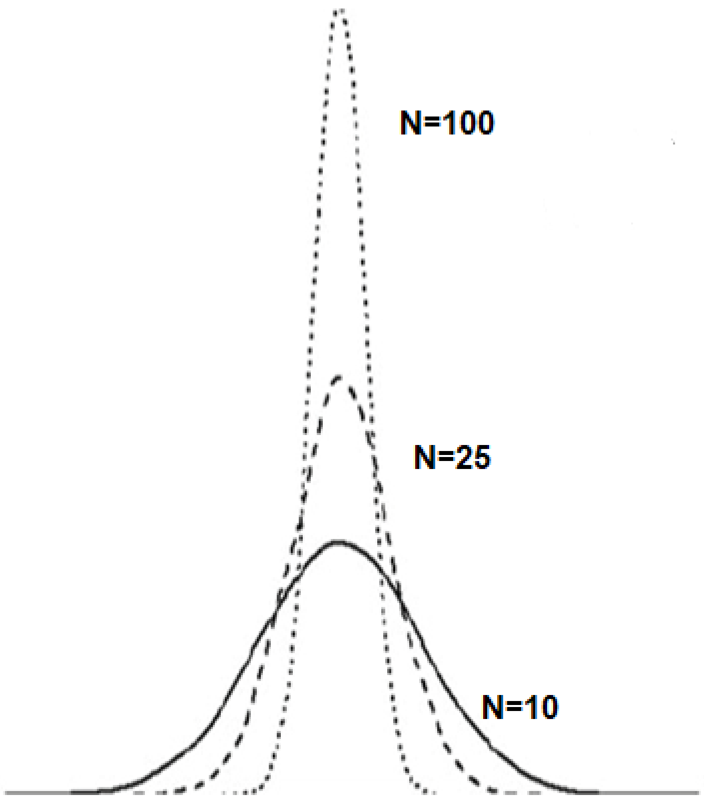
THE SAMPLING DISTRIBUTION

The mean of the sampling distribution is defined by our null hypothesis

In statistical tests, we assume that the mean of the sampling distribution is the same as the mean of the population distribution

However, we cannot just use the population standard deviation because the standard deviation is influenced by the number of observations we have in our sample

EXAMPLE: THREE SAMPLING DISTRIBUTIONS



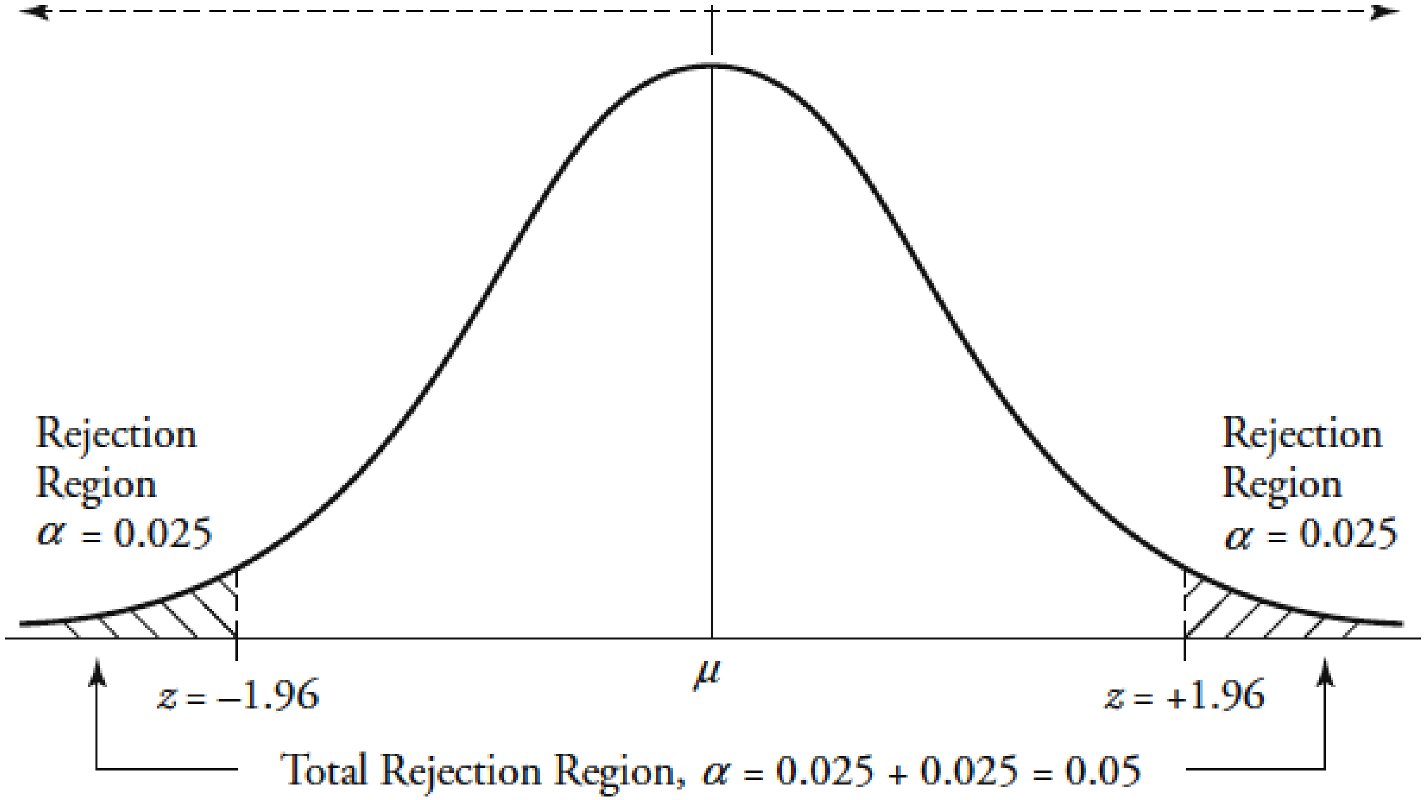
- The spread of the scores decreases as n increases
- To quantify this we rely on the standard error

$$se = \frac{\sigma}{\sqrt{N}}$$

EXAMPLE: FIND THE STANDARD ERROR OF THE SAMPLING DISTRIBUTION

$$\sigma_{sd} = \frac{\sigma}{\sqrt{N}} = \frac{15}{\sqrt{125}} = 1.342$$

REJECTION REGION FOR A .05 TWO-TAILED SIGNIFICANCE TEST



THE TEST STATISTIC

To calculate our test statistic, we can use the same formula we did in examining the relative position of a score, EXCEPT:

- We account for the fact that our *N* is larger and hence we are more precise; and
- We are now subtracting a MEAN and not a single observation…

Sample **mean**

$$z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{N}}}$$

Standard error

THE PROBLEM (AGAIN)

Let's say that you were interested in whether people who live in the West differ from people living in the US generally on average IQ scores.

The population characteristics for all people in the US are known. The mean score for the population is 100, and the standard deviation of the population mean is 15.

You conduct a study of 125 people living in California selected through an independent random sampling procedure from the population of all people who live in the state. You find that the mean IQ in your sample is 90.

$\mu = 100$

$\bar{X} = 90$

$\sigma = 15$

$N = 125$

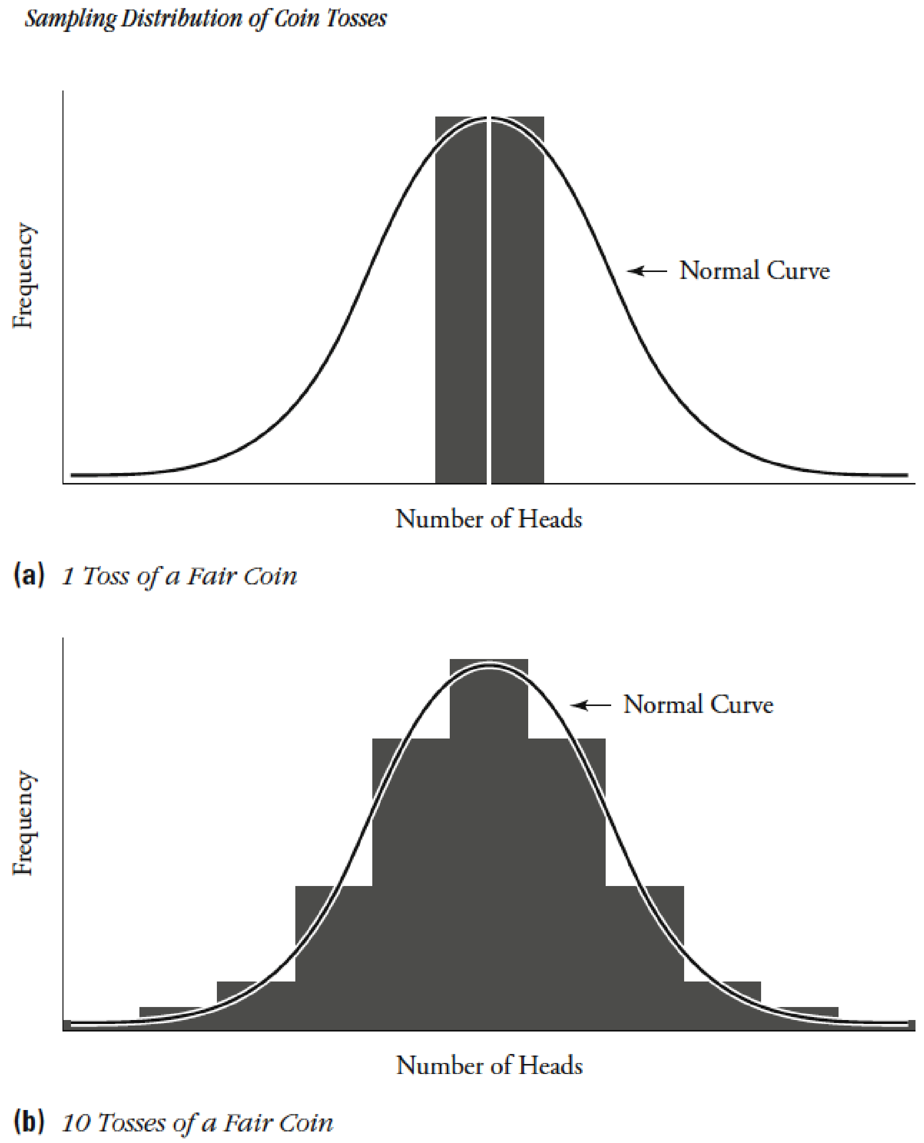
$$z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{N}}} = \frac{90 - 100}{\frac{15}{\sqrt{125}}} = -7.453$$

- Interpretation and Result?
- Because the test statistic is negative we are looking to the left of the curve
- If our calculated statistic is < -1.96 we reject the null hypothesis
- Since -7.45 < -1.96 we reject the null hypothesis
- The p-value is less than .05

APPLYING NORMAL SAMPLING DISTRIBUTIONS TO NON-NORMAL POPULATIONS

Often, things we study do not conform to normality

As we've seen before, however, many distributions that are non-normal in 1 sample will converge to normality as *N* gets big



THE CENTRAL LIMIT THEOREM

Sampling distributions begin to approximate the normal distribution as the sample size grows larger

More generally, the CLT says…

- If repeated independent random samples of size *N* are drawn from a population, then as *N* grows large, the sampling distribution of sample means will be approximately normal

Most statisticians believe that an N >= 50 will suffice

This means we can relax the assumption of normality, but ONLY when *N* is LARGE (and the other assumptions are satisfied)

THE SINGLE SAMPLE Z-TEST FOR PROPORTIONS

The central tendency of distribution that are based on categorical data (i.e. proportions) and its dispersion are measured by the mean and standard error, just as in distributions that develop from interval-level data.

Using the CLT, the sampling distribution for a proportion begins to approximate the normal in large samples

COMPUTING THE MEAN AND SD FOR A PROPORTION

It turns out that the mean of a proportion is equal to the proportion itself. But why?

Example: Five tosses of fair coin, Let X = 1 if heads results

Heads	Heads	Heads	Heads	Heads
1	1	0	0	1

What is the mean of this distribution? What is the probability of heads? They should be the same!

$$\mu = \mathbf{p}$$

CALCULATING THE STANDARD DEVIATION OF *P*

Same idea: take the sum of the squared deviations and divide by n-1 HOWEVER it turns out that there is an easier formula

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(X_i - \bar{X})^2}{N-1}} = \sqrt{p(1-p)}$$

Let’s get the standard deviation of our coin toss example…

INTUITION

When we state the proportion of successes expected under the null hypothesis, we also state by implication the mean and the standard deviation for the population distribution of scores

So if we state in the null hypothesis that the proportion of successes in the population is 0.50, we know that the mean of the population distribution of scores for our test of the null hypothesis is 0.50 and its variance is .25

$$\sigma = \sqrt{p(1-p)} = \sqrt{.50(.50)} = .50$$

TESTING HYPOTHESES FOR PROPORTIONS

Suppose that you were asked to evaluate a new education program. The foundation sponsoring the effort sought to achieve a program success rate of 75% among the 100,000 students enrolled in the program. Success was defined as completion of a six-month course supported by the foundation. Program personnel claim that the success rate is actually much greater than the criteria set by the foundation. However, a scathing critique of the program claims that the success rate of the program is actually much lower than 75%.

You are able to collect information on 150 students, selected using independent random sampling. You find that 85% of your sample successfully completed the course. What conclusions can you make, based on your sample results, about the claims of those haters?

ASSUMPTIONS

Level of measurement: interval – proportion

Population Distribution: Normal (*N is large)*

Sampling Method: Independent, random

Sampling Frame: 100,000 students enrolled in the program

Hypotheses:

H_0: The success rate of the program is .75 (i.e., $H_0: P = .75$

H_1: The success rate is not .75 (i.e., $H_1: P \neq .75)$

THE SAMPLING DISTRIBUTION

In calculating the mean and standard deviation or standard error for our sampling distribution, we rely on our null hypothesis.

Our null hypothesis states that the proportion of successes in the population is 75%. This means that the mean of the sampling distribution is also 0.75.

We need to calculate the standard error of the sampling distribution

We need to compute the z-score and select the rejection region

THE TEST STATISTIC

$$z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{N}}} = \frac{p - P}{\sqrt{P(1-P)}\Big/\sqrt{N}} = \frac{p - P}{\sqrt{P(1-P)/N}} = \frac{.85 - .75}{\sqrt{(.75)(.25)/150}} = 2.8329$$

Therefore, since 2.83 > 1.96 we reject the null. We know that $p < .05$ but can we get the exact p-value?

Note: this is the P under the null hypothesis NOT the *p* from the sample

COMPARING A SAMPLE TO AN UNKNOWN POPULATION: THE SINGLE SAMPLE T-TEST

The t-distribution has been proposed as an alternative sampling distribution when n is ‘small’ OR we do not have the population standard deviation (σ)

The t-distribution, like the chi-square, has different distributions, the shape is determined by the number of degrees of freedom (n – 1)

When the number of cases is > 500 the t – and z- distributions are identical, so use z

Typically used when n ~ 30 - 50 ish

The t distribution is provided in the ‘back of the book’ we will use sofware

EXAMPLE

Suppose that the study described earlier also examined the average test scores for those students who had completed the program. The foundation set a standard of success of 65 on the test. Again, folks believe that the average scores much higher than this while some believe its much lower. <u>Note: the question does not give you the population standard deviation!!</u>

In this case, you are able to take an independent random sample of **<u>51 students who have completed the test</u>**. You find that the test mean for the sample is 60, and the standard deviation is 15. What conclusions about the larger population of students can you come to based on your sample results?

$\mu = ?$

$\bar{X}$= ?

σ = ?

s= ?

N= ?

ASSUMPTIONS

Level of measurement

Population distribution

Sampling Method

Sampling Frame

Hypotheses $H_0: \mu = 65$ vs $H_1: \mu \neq 65$

The Sampling Distribution

THE TEST STATISTIC

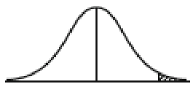
$$t = \frac{\bar{X}-\mu}{\sigma_{sd}} = \frac{\bar{X}-\mu}{\hat{\sigma}/\sqrt{N}} = \frac{\bar{X}-\mu}{\sqrt{\frac{\sum_{i=1}^{N}(X_i-\bar{X})^2}{N-1}}\Big/\sqrt{N}} = \frac{\bar{X}-\mu}{\sqrt{\frac{\sum_{i=1}^{N}(X_i-\bar{X})^2}{N}}\Big/\sqrt{N-1}} = \frac{\bar{X}-\mu}{s/\sqrt{N-1}}$$

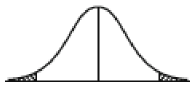
Applying the formula in this example gives

$$t = \frac{\bar{X}-\mu}{s/\sqrt{N-1}} = \frac{60-65}{15/\sqrt{51-1}} = \frac{-5}{2.1213} = -2.357$$

Be careful after this step: you need to look up the critical value for a two-sided test at the .05 level using the t-distribution, which here is -2.008. Since our calculated value is -2.008 we reject the null (i.e. it lies in our critical region)

Appendix 4 Critical Values of Student's *t* Distribution


One-tailed value


Two-tailed value

Degrees of Freedom	ONE-TAILED VALUE					
	0.25	0.10	0.05	0.025	0.01	0.005
	TWO-TAILED VALUE					
	0.50	0.20	0.10	0.05	0.02	0.01
1	1.000	3.078	6.314	12.706	31.821	63.657
2	0.816	1.886	2.920	4.303	6.965	9.925
3	0.765	1.638	2.353	3.182	4.541	5.841
4	0.741	1.533	2.132	2.776	3.747	4.604
5	0.727	1.476	2.015	2.571	3.365	4.032
6	0.718	1.440	1.943	2.447	3.143	3.707
7	0.711	1.415	1.895	2.365	2.998	3.499
8	0.706	1.397	1.860	2.306	2.896	3.355
9	0.703	1.383	1.833	2.262	2.821	3.250
10	0.700	1.372	1.812	2.228	2.764	3.169
11	0.697	1.363	1.796	2.201	2.718	3.106
12	0.695	1.356	1.782	2.179	2.681	3.055
13	0.694	1.350	1.771	2.160	2.650	3.012
14	0.692	1.345	1.761	2.145	2.626	2.977
15	0.691	1.341	1.753	2.131	2.602	2.947
16	0.690	1.337	1.746	2.120	2.583	2.921
17	0.689	1.333	1.740	2.110	2.567	2.898
18	0.688	1.330	1.734	2.101	2.552	2.878
19	0.688	1.328	1.729	2.093	2.539	2.861
20	0.687	1.325	1.725	2.086	2.528	2.845
21	0.686	1.323	1.721	2.080	2.518	2.831
22	0.686	1.321	1.717	2.074	2.508	2.819
23	0.685	1.319	1.714	2.069	2.500	2.807
24	0.685	1.318	1.711	2.064	2.492	2.797
25	0.684	1.316	1.708	2.060	2.485	2.787
26	0.684	1.315	1.706	2.056	2.479	2.779
27	0.684	1.314	1.703	2.052	2.473	2.771
28	0.683	1.313	1.701	2.048	2.467	2.763
29	0.683	1.311	1.699	2.045	2.462	2.756
30	0.683	1.310	1.697	2.042	2.457	2.750
31	0.682	1.309	1.696	2.040	2.453	2.744
32	0.682	1.309	1.694	2.037	2.449	2.739
33	0.682	1.308	1.692	2.035	2.445	2.733
34	0.682	1.307	1.691	2.032	2.441	2.728
35	0.682	1.306	1.690	2.030	2.438	2.724
40	0.681	1.303	1.684	2.021	2.423	2.704
45	0.680	1.301	1.680	2.014	2.412	2.690
50	0.680	1.299	1.676	2.008	2.403	2.678
55	0.679	1.297	1.673	2.004	2.396	2.669
60	0.679	1.296	1.671	2.000	2.390	2.660
70	0.678	1.294	1.667	1.994	2.381	2.648
80	0.678	1.293	1.665	1.989	2.374	2.638
90	0.678	1.291	1.662	1.986	2.368	2.631
100	0.677	1.290	1.661	1.982	2.364	2.625
120	0.677	1.289	1.658	1.980	2.358	2.617
>500	0.674	1.282	1.645	1.960	2.326	2.576

Since -2.3570 < -2.008 we reject the null

We will rely on program output