Chapter 8.

Single Sample Tests

Basic Concepts

- In this chapter, you are curious about whether there is significant difference between the mean of your sample and the mean of population, or between the proportion of your sample and the proportion of population.
- In this chapter, you are interested in applying this logic to one sample. Not interested in comparing two samples or more.
- Again, you will tackle the rest of your textbook in the similar way. REALLY IMPORTANT CHAPTER!

One sample t-test for the mean

- Null hypothesis: H0
- Alternative (or research) hypothesis: H1

1) If HO is true, this means you are ready to regard the difference as kind of coincidence! (your sample is really representative of population…)
2) If H1 is true, this means the result is not because of sampling error, rather there must be something that engenders such significant difference. (Too large to regard the difference as the result of error!)

The Statistics for Testing

- T-test… This means the statistics you are dealing with is “t”. (Student t…)
- What is t? You just use t (instead of z). Why? You don’t know about the standard deviation of population. Sometimes, your sample size is smaller than 50.
- $t = \frac{\bar{X} - \mu}{SE}$ Generally, $Statistic = \frac{(Observed - Null)}{S.E.}$
- Degree of freedom is n-1.

What does this equation imply?

- If your sample mean is much larger (or much smaller) than μ, what would happen to t? In this situation, can you insist that this is coincidence? Therefore, the chance of rejecting the null hypothesis becomes higher;
- If your mean is almost the same as μ, what would happen to t? In this case, can you make a big deal out of very small difference? Therefore, the chance of rejecting the null hypothesis becomes smaller.

Two types of errors in statistics

- Type 1: Even though null hypothesis is true, you are trying to reject that because you think the difference is big enough! Your guess entails type 1 error.
- Type 2: Even though null hypothesis is not right, you can't reject (not "accept", exactly saying) the hypothesis because you think the difference is not that big! Your estimation entails type 2 error.

- You already heard about “α”. This is none other than type 1 error.
- Recall what you learned at the last class. 95% confidence level means α is 0.05, which means there might be 5% error in your estimation of the mean (or proportion) of population.
- Even though the null hypothesis is true, that is, there is no big difference, the chance of making a mountain out of molehill is 5%!
- What if α=0.00? What does this mean?

Confidence interval, again

- You are dealing with t distribution. The shape of this distribution is different for different degree of freedom. Again, DF is n-1.
- N is 21. Your estimation falls somewhere around the population mean, but that might be smaller or bigger. (That is, two-tailed test) What is your confidence interval at the confidence level of 95%?

Critical (or Reject) Region

- Critical region means that your test statistics t falls in that region, you have to reject the null hypothesis because you can’t get away with such a big difference!
- H0: $X_{bar}=\mu$. What is H1? (Two-tailed test). Your sample mean is 10, but your population mean is 15. N is 21. SE=2.5.
- What is the critical region, then? Do you have to reject the null hypothesis, or not?
- You might as well look at P-value in the output. The smaller that value, the higher the chance of rejecting H0.

	alpha level for one-tailed test					
	0.1000	0.0500	0.0250	0.0100	0.0050	0.0005
	alpha level for two-tailed test					
df	0.2000	0.1000	0.0500	0.0200	0.0100	0.0010
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
6	1.440	1.943	2.447	3.143	3.707	5.959
7	1.415	1.895	2.365	2.998	3.499	5.405
8	1.397	1.860	2.306	2.896	3.355	5.041
9	1.383	1.833	2.262	2.821	3.250	4.781
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
more	1.282	1.645	1.960	2.326	2.576	3.29

One-tailed Test

- You already learned about two-tailed test, but sometimes you are curious about the direction of difference.
- If you are doing two-tailed test, you are not interested in the direction of difference. You are just wondering whether there is significant difference. If you are doing one-tailed test, you are more interested in, for example, whether your sample mean (or proportion) is bigger than the mean (or proportion) of population.

Why should we care about one or two tail?

- H0: $X_{bar}=\mu$. What is H1? Your sample mean is 10, but your population mean is 15. N is 21. SE=2.5. This was two-tailed test.
- What should H0 look like if your research hypothesis is that your sample mean is smaller than population mean?
- The highly negative your test statistic t, the chance of rejecting H0 becomes…
- What is the critical region?