Review of Inference

- ***What are your best estimates? Are you close?***
- ***Estimating the population mean, cont'd***
- ***Assume a Normal distribution***
- ***Confidence intervals I: Known variance of Y***
- ***… II: Unknown variance of Y***
- ***t distributions and standard errors***
- ***The t statistic***
- ***Critical values and symmetric confidence intervals***
- ***Hypothesis testing: Getting started***
- ***Rejection Rule I – t stats and critical values***
- ***Statistical Significance***
- ***p values***
- ***Rejection Rule II – p values and significance levels***
- ***Equivalence of Rejection Rules: An Example***

1. **What are your best estimates? Are you close?**
    a. **Estimation**: *What are your best estimates (of the unknown parameters)?*
        i. Covered in the Estimation review, where the emphasis was on linear and unbiased estimators (LUEs) and BLUE (minimum variance in the class of LUEs)
        ii. So far, we have made no assumptions about the specific distributions of random variables. We have just assumed they had well-defined means and variances, and that sampling was independent.
    b. **Inference**: *Are you sure/close?*
        i. Now we tackle Inference, and attempt to develop a sense of how close our parameter estimates are to the true underlying parameters. Unfortunately we can at best make only probabilistic statements. Our focus will be on the two main tools of inference: Confidence Intervals and Hypothesis Testing.
        ii. Confidence Intervals
            1. Confidence intervals provide an interval estimate of the true parameter value.
            2. Some high percent (95%?) of the interval estimates generated in some fashion (using an interval estimator, of course) will in fact contain the true unknown parameter value.
            3. But is the true parameter contained in the specific interval you are looking at? No idea!... though we do know that *some high percent (95%?) of the interval estimates generated in some fashion* … .
            4. And so we have some confidence about something… but what?

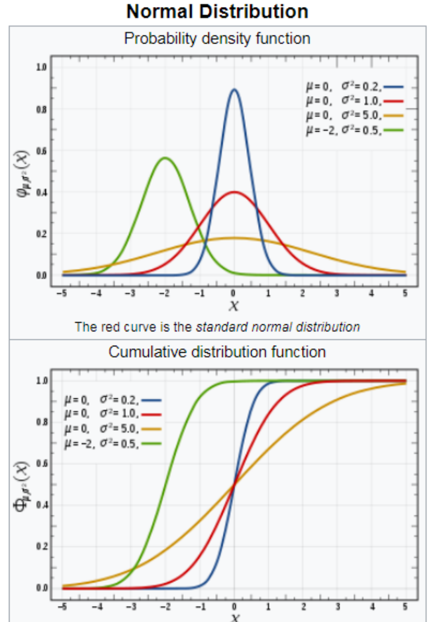
iii. Hypothesis Testing

1. Is the true (unknown) parameter zero? (other hypotheses can be tested… but overwhelmingly, this is the one we focus on)
2. Spoze that your point estimate is very very far from zero. It sure looks like the true parameter values isn't zero. But maybe you just had a really wacky unrepresentative sample, and the true underlying parameter was in fact zero. That is always a possibility… but is it probable? What's the probability that you would have seen the data/estimate that you saw being generated by a distribution with a zero parameter. Not high, you say? Well is it smaller than 10%? … than 5%? … than 1%?
3. If the probability of seeing what you saw when the true parameter was zero is less than say 5%, then we reject the Null Hypothesis that the true parameter is zero at the 5% *significance level*, and we say our estimate is *statistically significant* at that level. We could be wrong… but that is not at all likely!
4. The significance level is the maximum probability of making a judgement error of this sort (of concluding that that the true parameter is not zero, when in fact it is). And while we always have to allow for that possibility… we want the probability to be so so small… like less than 10%, … or 5%, … or even 1%.

iv. Distributional assumptions

1. To do *Inference*, construct *Confidence Intervals* and conduct *Hypothesis Tests*, we need to make assumptions about the specific distributions of the random variables that we are working with.
   a. This was not necessary for ***estimation***, where we made no distributional assumptions in showing that the Sample Mean was BLUE.
2. We typically assume ***Normal distributions***. You can of course work with other distributions… but you have to start somewhere, and why not begin with an assumption of Normality?

Normal Distribution



Notation	$\mathcal{N}(\mu, \sigma^2)$
Parameters	$\mu \in \mathbb{R}$ = mean (location) $\sigma^2 > 0$ = variance (squared scale)
Support	$x \in \mathbb{R}$
PDF	$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
CDF	$\frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right]$

As in our review of estimation, this review is built around what is probably the best known estimation example, estimating the mean of a distribution with randomly sampled data drawn from that unknown distribution.

The goal here is to keep things short and to focus on what is most important with regards to inference in SLR and MLR analysis. If you want more detail: Take a stats course!

2. **Estimating the Population Mean, cont'd:** Let's return to estimating the mean of the distribution of Y:

   a. As before, you have an iid random sample $\{Y_1, Y_2, \ldots Y_n\}$ from the distribution of Y, and you are using the sample mean, $\overline{Y} = \frac{1}{n}\sum Y_i$, to estimate the unknown mean of the distribution, μ.[1]

   b. You already know that $E(\overline{Y}) = \mu,\ and\ \ Var(\overline{Y}) = \frac{\sigma^2}{n}$, where σ^2 is the variance of Y. But since we've made no distributional assumptions yet, the particular nature of the distribution of Y (or of $\overline{Y} = \frac{1}{n}\sum Y_i$) is as yet unknown. But that will change shortly… as in ***Right Now!***

3. **Assume a Normal distribution**

   Assume that Y is Normally distributed, so that Y (and the Y_i's) are all $\mathrm{N}(\mu, \sigma^2)$.

   a. Since Y_i's are all $\mathrm{N}(\mu, \sigma^2)$ and the samples are independent, $\sum Y_i \sim \mathrm{N}(n\mu, n\sigma^2)$, and so

   $$\overline{Y} = \frac{1}{n}\sum Y_i \sim \mathrm{N}(\mu, \frac{\sigma^2}{n}) \ \ldots$$

   the sample mean is Normally distributed, with mean μ, variance $\frac{\sigma^2}{n}$, and standard deviation $sd = \frac{\sigma}{\sqrt{n}}$.[2]

   b. Put differently, $Z = \frac{\overline{Y} - \mu}{\sigma / \sqrt{n}} = \frac{\overline{Y} - \mu}{sd}$ has the standard Normal distribution. It's Normally distributed with mean 0 and variance 1… or as we sometimes say, $Z \sim N(0,1)$.

4. **Confidence Intervals I: *known variance*** σ^2

   a. We start by considering the case in which the variance of Y, σ^2, is known… but for some reason, its mean μ is unknown. This case is highly unrealistic (who would ever know the variance without knowing the mean?)… but it serves a useful pedagogic purpose.

   b. Further, we will focus on so called symmetric confidence intervals. They don't have to be symmetric, of course… but the symmetric case is easier to work through.

   c. Here's a symmetric (confidence) interval estimator: $\left[\overline{Y} - c\frac{\sigma}{\sqrt{n}}, \overline{Y} + c\frac{\sigma}{\sqrt{n}}\right]$, where $c \geq 0$ is some pre-specified *critical* value.

---

[1] Why use the Sample Mean? … well because it's a BLUE estimator, of course.

[2] Recall that sums of independent Normally distributed random variables will also be Normally distributed.

i. The only random component in this interval estimator is $\bar{Y}$, since n and the variance σ^2 are known, and c is pre-specified. The interval will shift around depending on $\bar{Y}$, the *Sample Mean* around which it is centered, and with a constant width of $2c\frac{\sigma}{\sqrt{n}}$.

In words: The Confidence Interval is the Sample Mean, a random variable, plus or minus c standard deviations (of $\bar{Y}$).

ii. The probability that this random interval estimator contains the unknown mean μ is: $prob\left(\mu \in \left[\bar{Y} - c\frac{\sigma}{\sqrt{n}}, \bar{Y} + c\frac{\sigma}{\sqrt{n}}\right]\right) = prob\left(\mu \in \left[\bar{Y} \pm c\frac{\sigma}{\sqrt{n}}\right]\right)$, which after some algebra can be re-expressed as $prob\left(-c \le \frac{\bar{Y} - \mu}{\sigma / \sqrt{n}} \le c\right)$.

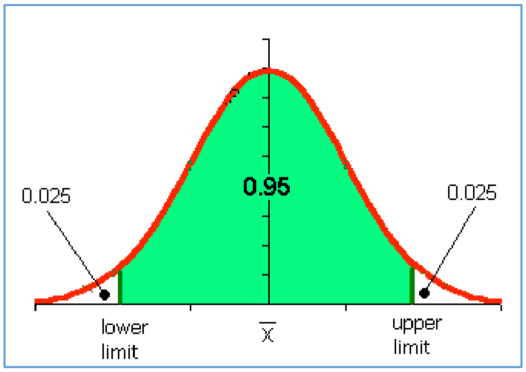
iii. But since $\frac{\bar{Y} - \mu}{\sigma / \sqrt{n}} = Z \sim N(0,1)$, this is just $prob\left(-c \le N(0,1) \le c\right)$. For a given level of confidence, this allows us to set the critical value c:

critical val. c	p(-c<Z<c)		p(-c<Z<c)	critical val. c
1.5	86.6%		89%	1.60
1.6	89.0%		90%	1.64
1.7	91.1%		91%	1.70
1.8	92.8%		92%	1.75
1.9	94.3%		93%	1.81
2	95.4%		94%	1.88
2.1	96.4%		95%	1.96
2.2	97.2%		96%	2.05
2.3	97.9%		97%	2.17
2.4	98.4%		98%	2.33
2.5	98.8%		99%	2.58

d. Some confidence levels and (symmetric) interval estimators:

i. 90% Confidence Interval: $\left[\bar{Y} \pm 1.64\frac{\sigma}{\sqrt{n}}\right]$

ii. 95% Confidence Interval: $\left[\bar{Y} \pm 1.96\frac{\sigma}{\sqrt{n}}\right]$



e. As you can see, a good rule of thumb is that the 95% confidence interval is the Sample Mean plus or minus two standard deviations.

5. **Confidence Intervals II:** ***unknown variance*** σ^2 (the much more realistic scenario)

   a. If the variance of Y is unknown, then we don't know the standard deviation of the estimator, $sd = \frac{\sigma}{\sqrt{n}}$. But we can compute the sample variance of the sample, $S_Y^2 = \frac{1}{n-1}\sum(Y_i - \bar{Y})^2$, which is an unbiased estimator of σ^2.

   b. And so $\frac{S_Y^2}{n}$ will be an unbiased estimator of the variance of the Sample Mean estimator: $Var(\bar{Y}) = \frac{\sigma^2}{n}$.

   c. Taking the square root, we have $\frac{S_Y}{\sqrt{n}}$ as an estimator of the standard deviation of $\bar{Y}$. We call $se = \frac{S_Y}{\sqrt{n}}$ the ***standard error*** (se) of the sample mean… it's an estimate of $sd = \frac{\sigma}{\sqrt{n}}$, the standard deviation of the (Sample Mean) estimator.

   d. ***Again***: The ***standard error*** is an approximation to the ***standard deviation*** of the Sample Mean estimator.

William Sealy Gosset

William Sealy Gosset (aka *Student*) in 1908 (age 32).

Born	13 June 1876 Canterbury, Kent, England
Died	16 October 1937 (aged 61) Beaconsfield, Buckinghamshire, England
Other names	Student
Alma mater	New College, Oxford
Known for	Student's t-distribution

6. **t Distributions and Standard Errors**

   a. If Y is normally distributed, then as before $\frac{\bar{Y}-\mu}{\sigma/\sqrt{n}} = \frac{\bar{Y}-\mu}{sd} \sim N(0,1)$, and replacing/estimating σ with S_Y, we have the estimator $\frac{\bar{Y}-\mu}{S_Y/\sqrt{n}}$.

   b. This estimator will have a t distribution with n-1 degrees of freedom. So $\frac{\bar{Y}-\mu}{S_Y/\sqrt{n}} \sim t_{n-1}$.

   c. The ***Student's t*** distribution was developed by William Sealy Gosset. in the early 1900's. At that time he was an employee (chemist and statistician) of Arthur Guinness & Son, the brewery in Dublin, Ireland. A brief bit of history from Wikipedia:[3]

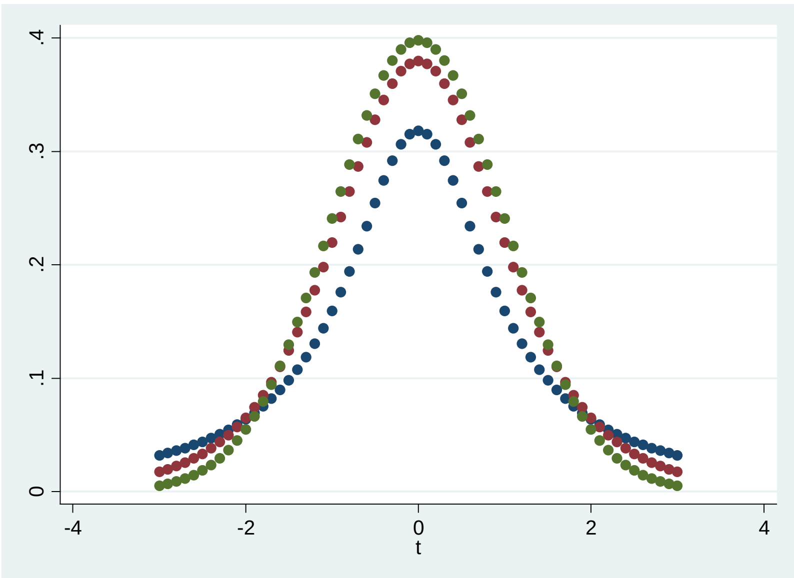
      > Gosset applied his statistical knowledge – both in the brewery and on the farm – to the selection of the best yielding varieties of barley. Gosset acquired that knowledge by study, by trial and error, and by spending two terms in 1906–1907 in the biometrical laboratory of Karl Pearson. Gosset and Pearson had a good relationship. …

---

[3] https://en.wikipedia.org/wiki/William_Sealy_Gosset

> Another researcher at Guinness had previously published a paper containing trade secrets of the Guinness brewery. To prevent further disclosure of confidential information, Guinness prohibited its employees from publishing any papers regardless of the contained information. However, after pleading with the brewery and explaining that his mathematical and philosophical conclusions were of no possible practical use to competing brewers, he was allowed to publish them, but under a pseudonym ("Student"), to avoid difficulties with the rest of the staff. Thus his most noteworthy achievement is now called Student's, rather than Gosset's, t-distribution. …
>
> It was, however, not Pearson but Ronald A. Fisher who appreciated the importance of Gosset's small-sample work, after Gosset had written to him to say I am sending you a copy of Student's Tables as you are the only man that's ever likely to use them!. Fisher believed that Gosset had effected a "logical revolution". Fisher introduced a new form of Student's statistic, denoted t. … The t-form was adopted because it fit in with Fisher's theory of degrees of freedom. Fisher was also responsible for applications of the t-distribution to regression analysis. …
>
> Gosset was a friend of both Pearson and Fisher, a noteworthy achievement, for each had a massive ego and a loathing for the other. He was a modest man who once cut short an admirer with this comment: "Fisher would have discovered it all anyway."

d. Here are the density functions for three t distributions, with dof's = 1, 5 and 99. Notice that the density function is symmetric and bell-shaped, and centered around 0. As the *dofs* increase, probability shifts from the tails to the middle of the distribution. In the limit, and as *dofs* approach infinity, the t distribution approaches $N(0,1)$.[4]



7. ***The t statistic***

a. $\dfrac{\overline{Y}-\mu}{S_Y/\sqrt{n}}$ is sometimes called the ***t-statistic***, and it drives inference (when using the Sample Mean to estimate the unknown mean, and the variance is unknown).

b. ***Worth repeating!*** The t statistic drives inference!... and, Surprise!... it has a t distribution with n-1 dofs.

---

[4] Sometimes we say that the t-distribution has *fatter tails* than the Standard Normal distribution.

8. Critical values and symmetric confidence intervals

a. To derive Confidence Intervals for the case of now unknown variance σ^2, we proceed as above… the only difference being that we are now working with the t distribution (rather than the Normal distribution), and working with the standard error of $\bar{Y}$ (rather than its now unknown standard deviation).

b. Here's a symmetric (confidence) interval estimator (the Sample Mean plus or minus c standard errors):

$$\left[\bar{Y} - c\frac{S_Y}{\sqrt{n}}, \bar{Y} + c\frac{S_Y}{\sqrt{n}}\right] \text{ or } \left[\bar{Y} \pm c\frac{S_Y}{\sqrt{n}}\right]$$

where $c \geq 0$ is some pre-specified "critical" value, determined, as described below, using the t distribution with n-1 dofs.

i. As before, and after some algebra,

$$prob\left(\mu \in \left[\bar{Y} - c\frac{S_Y}{\sqrt{n}}, \bar{Y} + c\frac{S_Y}{\sqrt{n}}\right]\right) = prob\left(-c \leq \frac{\bar{Y} - \mu}{\sigma / \sqrt{n}} \leq c\right) = prob\left(-c \leq t_{n-1} \leq c\right),$$

where t_{n-1} is a t distribution with n-1 dofs.

dofs	90.0%	92.5%	95.0%	97.5%	99.0%
5	2.02	2.24	2.57	3.16	4.03
10	1.81	1.99	2.23	2.63	3.17
15	1.75	1.91	2.13	2.49	2.95
20	1.72	1.88	2.09	2.42	2.85
25	1.71	1.86	2.06	2.38	2.79
30	1.70	1.84	2.04	2.36	2.75
50	1.68	1.82	2.01	2.31	2.68
75	1.67	1.81	1.99	2.29	2.64
100	1.66	1.80	1.98	2.28	2.63
infinite	***1.64***	***1.78***	***1.96***	***2.24***	***2.58***

ii. For a given level of confidence, and given numbers of dofs, we set the critical value c using the t_{n-1} distribution.

iii. 90% Confidence Interval

1. Degrees of freedom - 25: $\left[\bar{Y} \pm 1.71\frac{S_Y}{\sqrt{n}}\right]$

2. Degrees of freedom – infinite, N(0,1): $\left[\bar{Y} \pm 1.64\frac{S_Y}{\sqrt{n}}\right]$

iv. 95% Confidence Interval

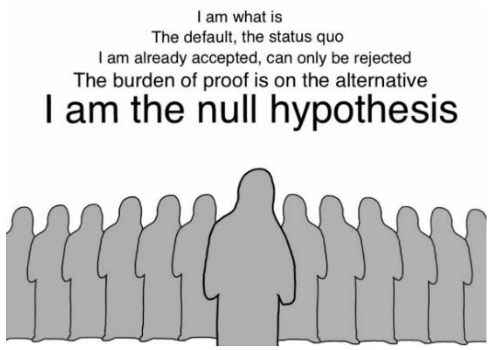
1. Degrees of freedom - 25: $\left[\bar{Y} \pm 2.06\frac{S_Y}{\sqrt{n}}\right]$

2. Degrees of freedom – infinite, N(0,1): $\left[\bar{Y} \pm 1.96\frac{S_Y}{\sqrt{n}}\right]$

c. A good rule of thumb: For a 95% confidence interval, use the sample mean $\pm$ a couple standard errors. It's not going to be precisely correct, but it's an easy and altogether not-that-bad approximation. And you can use the same rule when you get to SLR and MLR analyses.

9. **Hypothesis testing: Getting started**

   a. **Null and Alternative Hypotheses**

   - H_0: the *Null* Hypothesis (the hypothesis we are testing)
   - H_1: the *Alternative* Hypothesis, the alternative to H_0

   b. Overwhelmingly, the Null Hypothesis we will be focusing on is:

   H_0: The true parameter value is 0

   c. Two types of Error:

   - **Type I – False Rejection**:

     Rejecting the Null, H_0, (and accepting the Alternative, H_1) when H_0 is true

   - **Type II – False Acceptance**:

     Accepting the Null, H_0, (and rejecting the Alternative, H_1) when H_0 is false

   d. We generally focus on Type I error and *protect* the Null hypothesis… only rejecting the Null hypothesis in the face of *overwhelming* evidence to the contrary… where, for example, the probability of being wrong (incorrectly rejecting the Null) is very small… $\leq$ 10%, … $\leq$ 5%, or … $\leq$ 1%, or… . So while mistakes may happen, their probability will be small small small.

   e. ***Significance levels*** (α):

   We call these probabilities *significance levels*, and typically denote them with α. They are the maximum acceptable probability of a Type I error = P(Reject H_0 | H_0 is true)

   i. Where do significance levels come from? *We make them up!*

10. Continuing with our example: Sampling from a Normal distribution with unknown variance, and testing $H_0 : \mu = 0$ (far and away the most common hypothesis test).

    a. As before, we randomly sample n times (iid) from $Y \sim N(\mu, \sigma^2)$, and use the Sample Mean ($\bar{Y}$) to estimate the true mean, μ. Why the Sample Mean? Well, because it's BLUE, of course.

    b. Consider again the ***t-statistic***: $\dfrac{\bar{Y} - \mu}{S_Y / \sqrt{n}}$. The numerator, $\bar{Y} - \mu$, tells you how far the sample mean, $\bar{Y}$, is from the true mean, μ. Dividing by $se = \dfrac{S_Y}{\sqrt{n}}$, expresses that difference in units of standard errors, which provides for greater comparability. So the t statistic tells you how many standard errors the sample mean is from the true mean μ.

    i. Note that the sign of the t-statistic will depend on the estimated mean, so t statistics can be positive or negative.

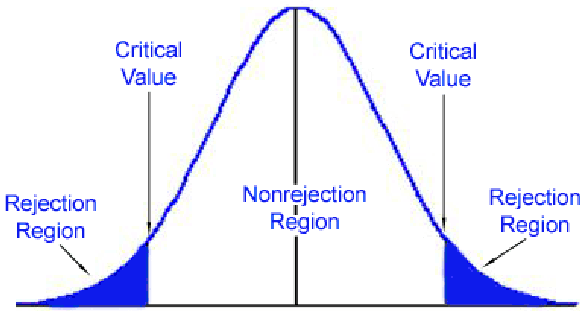
c. Under the Null Hypothesis ($H_0 : \mu = 0$), $\mu = 0$ and the t statistic, $t\ stat = \frac{\bar{Y}}{S_Y / \sqrt{n}}$, has a t distribution with n-1 degrees of freedom. As before, we express this as $\frac{\bar{Y}}{S_Y / \sqrt{n}} \sim t_{n-1}$.

i. This is the ***t-stat*** under the Null Hypothesis ($H_0 : \mu = 0$). (Notice that I did not call it a ***t statistic***… if you see or hear the term ***t stat***, unless you know otherwise, it's reasonable that there's an underlying assumption that the true parameter value is zero)

11. **Rejection Rule I – t stats and critical values**:

a. Reject if $\left|\frac{\bar{Y}}{S_Y / \sqrt{n}}\right| > c$ … so reject if the t stat is larger in magnitude than some *critical value* $c > 0$… Or in other words: reject the Null Hypothesis that the true mean is 0 if $\bar{Y}$ is more than c standard errors away from that value, 0.

b. So reject the Null Hypothesis $H_0 : \mu = 0$ if $\frac{\bar{Y}}{S_Y / \sqrt{n}} > c$ or if $\frac{\bar{Y}}{S_Y / \sqrt{n}} < -c$.



i. The rejection region consists of two *tails* of the t-distribution… which is why this is called a **two-tailed** test.

c. The probability of Type I error is the probability of rejecting H_0 when it is true. For this test, that probability is $prob\left(\left|\frac{\bar{Y}}{S_Y / \sqrt{n}}\right| > c\right)$.

d. But since $\frac{\bar{Y}}{S_Y / \sqrt{n}} \sim t_{n-1}$ under the Null Hypothesis,

$prob\left(\left|\frac{\bar{Y}}{S_Y / \sqrt{n}}\right| > c\right) = prob\left(\left|t_{n-1}\right| > c\right)$ is just the probability that we're in the tails of a t distribution with n-1 degrees of freedom (below $-c$ or above $+c$).

12. Suppose that we want to select the critical value c so that the probability of falsely rejecting the Null hypothesis is $\alpha = 5\%$ (the *significance level* of the test). Then because the t-distribution is symmetric, we want to find c^* such that:

$$prob\left(\frac{\overline{Y}}{S_Y / \sqrt{n}} > c^*\right) = prob\left(t_{n-1} > c^*\right) = .025$$ (focusing just on the upper tail probability).

Critical Values for the t and Standard Normal distributions

Degrees of Freedom	***Significance Levels (two-tailed test)***			
	20%	***10%***	***5%***	***1%***
5	1.48	2.02	2.57	4.03
10	1.37	1.81	2.23	3.17
15	1.34	1.75	2.13	2.95
20	1.33	1.72	2.09	2.85
25	1.32	1.71	2.06	2.79
30	1.31	1.7	2.04	2.75
35	1.31	1.69	2.03	2.72
40	1.3	1.68	2.02	2.7
45	1.3	1.68	2.01	2.69
50	1.3	1.68	2.01	2.68
N(0,1)	1.28	1.64	1.96	2.58

a. For this critical value, c^*, we then reject $H_0 : \mu = 0$ if $|t\ stat| > c^*$ or, equivalently, if $|\overline{Y}| > c^* \frac{S_Y}{\sqrt{n}}$.

b. With this test, we will falsely reject H_0 5% of the time… a low Type I error rate (a small risk that we rejected the Null Hypothesis when it was true).

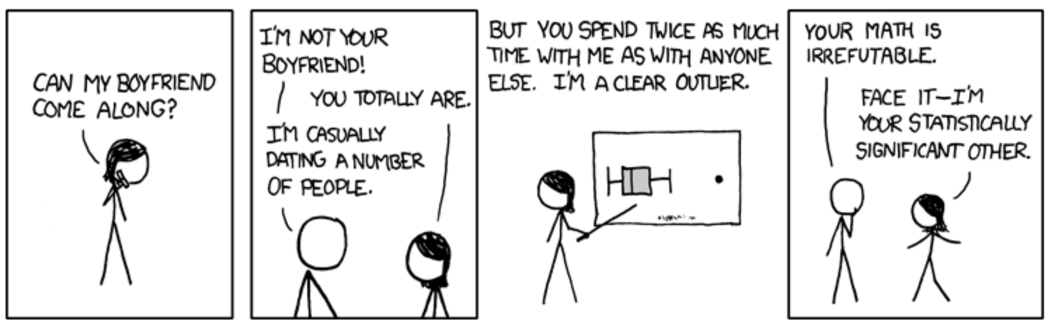
c. So the test is: Reject the Null Hypothesis if …

   - the observed sample mean is at least c^* Standard Errors away from 0, or put differently,
   - if the t-stat is larger in magnitude than c^*,

   where the particular value of c^* reflects the significance level of the test, α , and the degrees of freedom.

d. ***Statistical Significance***

- If we can reject the Null Hypothesis $H_0 : \mu = 0$ at, say, the 5% significance level, then we say that the estimate is ***statistically significant*** at the 5% level (using a two-tailed test).
- It makes no sense to talk about statistical significance without referencing the significance level. Common significance levels are 10%, 5%, 1% and .1%. Where do they come from? *We make them up!*
- And remember: Every estimate is statistically significant at some significance level! … but in some cases, that level is embarrassingly large!



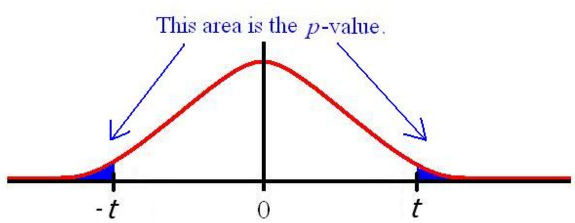
13. Probability Values … or *p values*

a. ***Probability values*** *(or p* values for short) are the maximum significance levels at which we can conduct a hypothesis test and fail to reject the Null Hypothesis…. Typically, p values are reported for two-tailed tests.

   i. So if the p value is, say .02, then the null hypothesis can be rejected at significance levels above 2%, but not at smaller significance levels.

b. More formally: Suppose we have a particular sample mean $\bar{y}$ and particular standard error $se = \frac{S_y}{\sqrt{n}}$. Then we can determine the p value as the probability of being that far or further away from 0 under the Null Hypothesis that $\mu = 0$:

$prob\left(\left|t_{n-1}\right| > \left|t\ stat\ \right|\right) = p$.



14. **Rejection Rule II – p values and significance levels**

a. We will reject the two-tailed Null Hypothesis that $\mu = 0$ at the significance level α if and only if $p < \alpha$ (the p-value for the given sample is smaller than the significance level for the test):

i. From before, we reject the Null hypothesis if the t stat is larger than the critical value. So reject if:

$$prob\left(\left|\frac{\bar{Y}}{S_Y/\sqrt{n}}\right|\right) > c_\alpha \text{ where } c_\alpha \text{ is defined by } prob\left(\left|t_{n-1}\right| > c_\alpha\right) = \alpha .$$

ii. But the t stat is larger than the critical value if and only if the p value is less than the significance level, so we can equivalently reject if $p < \alpha$.

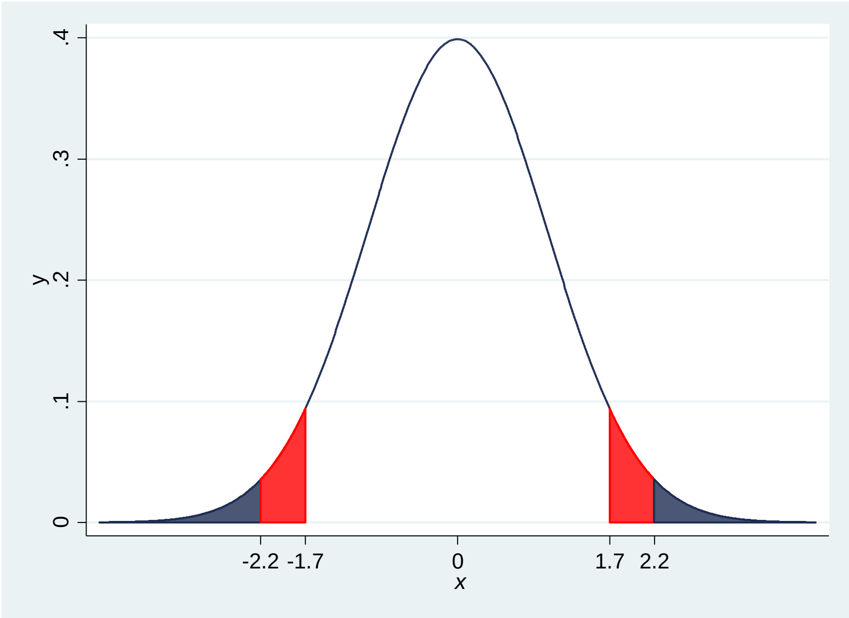
iii. Since ***small p values ~ large t stats***, we reject if: $p < \alpha \Leftrightarrow \left|\frac{\bar{y}}{S_y/\sqrt{n}}\right| = \left|t\ stat\right| > c_\alpha$

15. Equivalence of Rejection Rules: Example

a. Here's an example:

i. In this case $\alpha = .10$, $dofs = 30$ *and* $c^* = 1.7$. So we reject the Null Hypothesis that $\mu = 0$ if $\left|t\ stat\right| > 1.7$ or if $p\ value < .10$.

ii. Here, $\left|t\ stat\right| = 2.2 > 1.7$ and $p\ value < .10$, since the shaded region to the right of 2.2, $\frac{p\ value}{2}$, is less than the shaded region to the right of 1.7, $\frac{\alpha}{2}$. So: *Reject! Reject!*



b. Accordingly: p values make hypothesis testing easier, since we don't need to determine critical values. If we want the Type I Error to be less than 10% in a two-tailed test, then we reject the null hypothesis only if the p value is below 10%. Done!

c. You'll discover that Stata gives you the p-values in the regression output… making hypothesis testing and the determination of statistical significance a snap!