Estimation of Parameters from Statistics:

Sample--> Measure--> Statistic=====> Parameter

Point estimates:

Sample Mean, Xbar => μ

Sample S => σ

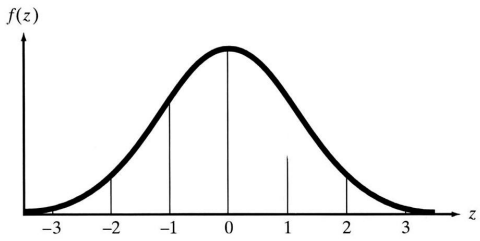
No statement on precision or if it might be "close" to the "true value".

Confidence Interval:

Consider the distribution of Xbar - μ for numerous samples. 95% of all estimates will fall within +/- 1.96 standard errors of the true mean, μ.
Thus, if we consider Xbar +/- 1.96 SEM, we would be 95% confident that this interval includes the true mean, μ.

If the standard deviation, σ, isn't known, then it is estimated from the sample S, and we use the t statistic instead of Z.

6 σ limits (3 on either side of the mean) include 99.73% of all values. For practical purposes, this is considered the "range" of a process output and describes "process capability." The goal is to have a process which will produce acceptable product consistently, or with "6 σ limits"



Comparing two independent Samples

Distribution of the differences:

Mean of $X_1 - X_2 = \mu_1 - \mu_2$.

Standard error of the difference is a function of the two σ s:

$$SE(X_1 - X_2) = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

If we assume that true σ isn't known and that the two samples are estimates of a common population σ, then the pooled estimate of variance is a weighted mean:

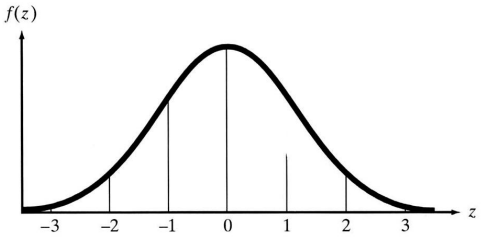
$$S_p^2 = \frac{(n1-1)s_1^2 + (n2-1)s_2^2}{n1+n2-2}$$

Confidence interval for an estimate of the difference between means of the two populations uses this common estimate of s

$$S = \sqrt{s_p^2(\frac{1}{n_1} + \frac{1}{n_2})}$$

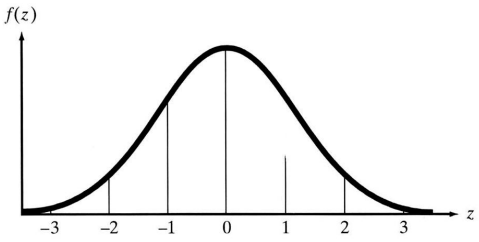
Paired Observations: The "Before & After" Experiment

- When observations aren't independent, pairs may act as controls for all sorts of unknown variable factors.
- By pairing observations and only working with the distribution of differences as single observations, we lose degrees of freedom, but eliminate all sorts of other individual variation.
- Test for deviation of the mean of the differences from zero.



How to get tighter confidence limits

1. Change measurement technique to reduce random variation
2. Reduce the level of confidence,
   e.g. be happy with 95% confidence (Z= 1.96)
   rather than 99% confidence (Z= 2.58).
3. Increase the sample size to tighten up the SEM.



How large a sample?

1. State confidence level desired, find the corresponding Z or t value.
2. Estimate σ_X.
3. State the difference you want to detect.
4. Solve for n.

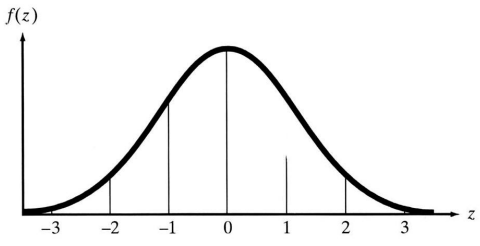
$$Z = \frac{X-\mu}{\frac{\sigma}{\sqrt{n}}}$$

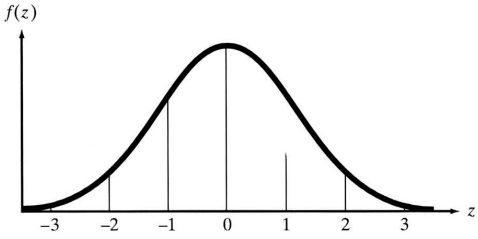
Example 8.11:

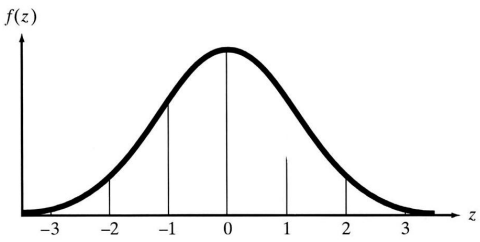
- want 99% confidence
- s = 1.6 day
- want within 0.5 day

Tests of "Statistical Significance"

1. State Null Hypothesis: H_0 , that population mean = hypothesized value, $\mu = \mu_0$. i.e. there's no difference.
   - H_0: is the boring status quo that we assume unless someone presents overwhelming evidence to the contrary.
     - Usually something like "the suspect is innocent" or "there is no difference"
   - H_A (sometimes called H_1) is the more interesting result that is speculated about.
     - This "proves" that our product is better than the other guy's!
     - Uh oH! production ***is*** all screwed up!
2. Choose significance level, α.
3. Compute the test statistic for the SEM, e.g. Z or t score
4. Determine the critical region, with α in one tail or $\alpha/2$ in each of two tails.
5. Reject H_0 if $Z_{observed} > Z_{critical}$







Tests of "Statistical Significance"

The meaning is that:

1. if Z is in the non-reject region: "There is a ___% chance that this difference could happen just by chance. There is little reason to think that this difference implies that the difference is real. It could just be due to sampling variation."

2. If Z is in the reject tail: "The difference could still be due to sampling variation, but there is less than a ___% probability that the means are the same and this difference would occur just by chance. I can conclude that the means may really be different."

Conventions in literature tables:

$P > 0.05$ (non-reject region), "result not significantly different"
$P < 0.05$ (reject at S.L.= .05), "significant", one star
$P < 0.01$ (reject at S.L.= .01), "Highly significant", two stars
$P < 0.005$ or 0.001 Investigator highly delighted, many stars.

Hypothesis testing: Two ways to be wrong

	State of nature	
Conclusion	H_0 true	H_0 false
Accept H_0		Type II, β
reject H_0	Type I, α	Power $1-\beta$