Stat 10, Section 1A
Thursday, May 13th, 2010

David Diez

1 Reminder: lab summaries has word and sentence limits

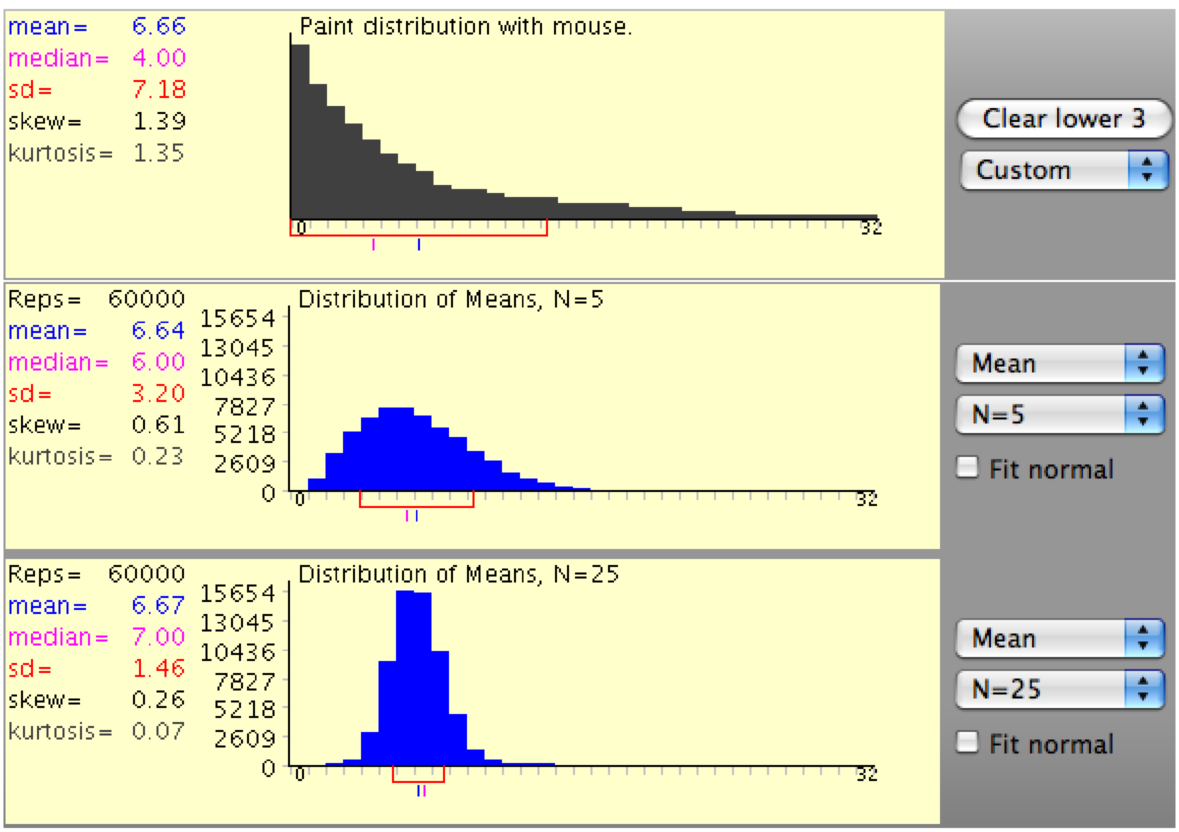
Each part should be only **one sentence**. For all parts there should be a total of **no more than 150 words**.

2 Lab 6 (pennies) comments

Sample distribution: The distribution of the individual observations from a sample.
Sampling distribution: The distribution of a statistic, usually the sample mean.

The purpose of this lab was to study the behavior of the sampling distribution of the mean as a response to sample size and distribution shape.

There were also a lot of conclusions that had little to do with the listed purpose of the lab.



Identifying the Proper Confidence Interval or Test

When creating a confidence interval (CI) or running a test for a mean, proportion, or difference in means or proportions, there is some *estimate* of the mean, proportion, or difference from the data. There is also the standard error (SE), which is a measure of the uncertainty of this *estimate*. The general form for a $100 * (1 - \alpha)\%$ confidence interval is

$$estimate \pm (critical\ value) * SE_{estimate}$$

The *critical value* is found from the appropriate table (normal or t table). ($\alpha = 0.05$ for a 95% confidence interval.) The general form for a hypothesis test statistic is

$$test\ statistic = \frac{estimate - parameter}{SE_{estimate}}$$

where *parameter* is the value under question in H_0. For instance, if $H_0 : \mu_1 - \mu_2 = 7.3$, then $parameter = 7.3$. Alternatively, if $H_0 : p = 0.3$, then $parameter = 0.3$.

To identify the *estimate* and $SE_{estimate}$, begin by asking

How many samples are there?	1	2
Are proportions or means of interest?	proportion	mean
Is this for a confidence interval or a hypothesis test?	CI	test

Using these answers, identify the proper CI or test in the table. Additional instructions and special circumstances for using the table below:

- If the data is for proportions or the standard deviation is known, use Z (ie, the normal dist.). If it is for means and the standard deviation is unknown, use t (ie, the t-dist.).
- If you chose (2, proportion, test) from above, only use the pooled test if H_0 is $p_1 - p_2 = 0$ (or $p_1 = p_2$). For the pooled test, $\hat{p}_c = \frac{x_1 + x_2}{n_1 + n_2} = \frac{n_1 \hat{p}_1 + n_2 \hat{p}_2}{n_1 + n_2}$.
- If you chose (2, mean, either CI or test) and the data is paired, work only with the differences. n_{diff} = # of differences (ie, 2 samples each of size 10 implies there are 10 differences, $n_{\text{diff}} = 10$).
- To avoid any confusion: s, s_1, s_2, and s_{diff} are standard deviations of the samples.

Circumstance	parameter	estimate	$SE_{estimate}$
1-prop (CI)	p	$\hat{p}$	$\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$
1-prop (test, $p_0 = expected$)	p	$\hat{p}$	$\sqrt{\frac{p_0(1-p_0)}{n}}$
2-prop (unpooled, test or CI)	$p_1 - p_2$	$\hat{p}_1 - \hat{p}_2$	$\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$
2-prop (pooled, test only)	$p_1 - p_2$	$\hat{p}_1 - \hat{p}_2$	$\sqrt{\frac{\hat{p}_c(1-\hat{p}_c)}{n_1} + \frac{\hat{p}_c(1-\hat{p}_c)}{n_2}}$
1-samp (t-test or CI)	μ	$\bar{x}$	$s/\sqrt{n}$
2-samp (unpaired, t-test or CI)	$\mu_1 - \mu_2$	$\bar{x}_1 - \bar{x}_2$	$\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$
2-samp (paired, t-test or CI)	$\mu_{\text{diff}} = \mu_1 - \mu_2$	$\bar{x}_{\text{diff}}$	$s_{\text{diff}}/\sqrt{n_{\text{diff}}}$