**STAT 6200 — Design of Experiments for Research Workers**
**Lab 3 – Due Tuesday, October 19**

**Example — Body Weights for Diabetics:**

Percentages of ideal body weight were determined for 18 randomly selected insulin-dependent diabetics and are shown below. A percentage of 120 here means that an individual weighs 20% more than his or her ideal body weight; a percentage of 95% mean that the individual weighs 5% less than the ideal.

$$107, 119, 99, 114, 120, 104, 88, 114, 124, 116, 101, 121, 152, 100, 125, 114, 95, 117$$

These data are also contained in the Minitab worksheet DIABETIC.MTW on the course web site. Retrieve this file, open it in Minitab, and follow along.

1. Suppose that it is known that percentage of ideal body weight is normally distributed with standard deviation of $\sigma = 10$ in this population. Find a 95% confidence interval for $\mu$, the mean percentage of ideal body weight among insulin-dependent diabetics.

**Answer:** Let $x_i$ =the percentage of ideal body weight for subject $i$, $i = 1, \ldots, 18$. That is, $x_1 = 107, x_2 = 119, \ldots, x_{18} = 117$. Then the sample mean is

$$\bar{x} = \frac{1}{18} \sum_{i=1}^{18} x_i = \frac{1}{18}(107 + \cdots + 117) = 112.78.$$

The formula for a $100(1-\alpha)\%$ CI for $\mu$ when $\sigma$ is known is

$$\bar{x} \pm z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \qquad (*)$$

For a 95% interval, $100(1-\alpha)\% = 95\%$ which means that $\alpha = .05$. Therefore, we need $z_{1-\alpha/2} = z_{1-.05/2} = z_{.975}$. We can find this value either by

a. looking up $1 - .975$ in Table A.3 in the back of our book, or
b. using Minitab.

In Minitab 14, the commands are

Calc → Probability Distributions → Normal...

Then select "Inverse Cumulative Probability", set "Mean"=0.0, set "Standard deviation"=1.0, select "Input constant", enter .975 for the constant, and then hit OK.

Either way, we find that $z_{.975} = 1.96$.

So, a 95% CI for $\mu$ is

$$\begin{aligned}\bar{x} \pm z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} &= 112.78 \pm 1.96\frac{10}{\sqrt{18}} \\ &= 112.78 \pm 1.96\underbrace{(2.357)}_{=\text{the s.e. of } \bar{x}} \\ &= (112.78 - 1.96(2.357), 112.78 + 1.96(2.357)) \\ &= (108.16, 117.40)\end{aligned}$$

These calculations can be done more automatically in Minitab 14 as follows: select

Stat → Basic Statistics → 1-Sample Z...

Then under "Samples in columns:" select the variable pctidealwgt, enter 10 in the box next to "Standard deviation:", and hit OK.

These steps give a 95% CI of (108.16, 117.40) which agrees with the results we obtained by hand.

- Notice also that along with the confidence interval, Minitab outputs the sample mean $\bar{x} = 112.778$, the sample standard deviation $s = 14.424$, and the standard error of the mean s.e.$(\bar{x}) = 2.357$, which agrees with the result we computed above.

2. Assuming the same circumstances as in part 1, compute a 99% confidence interval for $\mu$ rather than a 95% interval.

**Answer:** A $100(1-\alpha)\%$ interval is still given by formula (*). Now, however, $100(1-\alpha) = 99$, so that $\alpha = .01$. Therefore, we now need $z_{1-\alpha/2} = z_{1-.01/2} = z_{.995}$. This quantity can be obtained from Table A.3, but it is easire at this point to use Minitab. Again, select

Calc → Probability Distributions → Normal...

Then select "Inverse Cumulative Probability", set "Mean"=0.0, set "Standard deviation"=1.0, and select "Input constant". However, now enter .995 for the constant, and then hit OK. The answer is $z_{.995} = 2.57583$. So, now our 99% CI for $\mu$ is

$$\begin{aligned}\bar{x} \pm z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} &= 112.78 \pm 2.57583\frac{10}{\sqrt{18}} \\ &= 112.78 \pm 2.57583(2.357) \\ &= (112.78 - 2.57583(2.357), 112.78 + 2.57583(2.357)) \\ &= (106.71, 118.85)\end{aligned}$$

Alternatively, we can get Minitab to do the work exactly as before: select

Stat $\rightarrow$ Basic Statistics $\rightarrow$ 1-Sample Z...

Then under "Samples in columns:" select the variable pctidealwgt, and enter 10 in the box next to "Standard deviation:". But before hitting OK, go to "Options...", set "Confidence level:" to 99, and then hit OK twice. The answer is $(106.71, 118.85)$ which agrees with what we computed by hand.

- Notice that the confidence interval is wider for a higher confidence level.

3. Suppose now the population standard deviation is unknown and must be estimated from the sample data. Recompute the 99% interval from part 2.

**Answer:** For a sample of size $n = 18$ from a normal distribution, we can form our confidence interval using a $t$ distribution rather than a $Z$, or standard normal.

If $t_p(d)$ denotes the $100p$th percentile of a $t$ distribution with $d$ degrees of freedom, then a $100(1-\alpha)\%$ CI for $\mu$ is given by

$$\bar{x} \pm t_{1-\alpha/2}(n-1)\frac{s}{\sqrt{n}} = \bar{x} \pm t_{1-\alpha/2}(n-1)\text{s.e.}(\bar{x}). \qquad (**)$$

- Notice that the only difference between the formula for the $\sigma$ known case that we had in part 2, and the $\sigma$ unknown case that we have here, is that we replace $z_{1-\alpha/2}$ with $t_{1-\alpha/2}(n-1)$, which is a bigger quantity.
  - The idea here is that having to estimate $\sigma$ by $s$ introduces additional undertainty into the problem of estimating $\mu$. Therefore, we go plus or minus $t_{1-\alpha/2}(n-1)$ standaard errors rather than plus or minus $z_{1-\alpha/2}$ standard errors, which makes our interval wider to reflect the greater uncertainty in this situation.

Based on the original data, we can compute the sample standard deviation as

$$\sqrt{\frac{1}{n-1}\{(\sum_{i=1}^{n} x_i^2) - n\bar{x}^2\}} = \sqrt{\frac{1}{18-1}\{(107^2 + \cdots + 117^2) - 18(112.778^2)\}}$$
$$= 14.424$$

Now we want a 99% confidence interval, which corresponds to $\alpha = .01$, so we need $t_{1-\alpha/2}(n-1) = t_{1-.01/2}(18-1) = t_{.995}(17)$. We can get this quantity from Table A.4 in the back of the book: $t_{.995}(17) = 2.898$. Alternatively, we can get it from Minitab by selecting

Calc $\rightarrow$ Probability Distributions $\rightarrow$ t...

Then select "Inverse Cumulative Probability", set "Degress of freedom:"=17, select "Input constant", enter .995 for the constant, and then hit OK. Minitab gives the slightly more accurate answer $t_{.995}(17) = 2.89823$.

- Notice that this $t$ value, 2.89823, is slightly larger than the corresponding $z$ value from part 2, 2.57583, when $\sigma$ was known. That means that here the 99% confidence interval will be wider than the corresponding interval in part 2.

Plugging the values we've computed into formula (**) we get

$$\begin{aligned} \bar{x} \pm t_{1-\alpha/2}(n-1)\frac{s}{\sqrt{n}} &= 112.778 \pm 2.89823\frac{14.424}{\sqrt{18}} \\ &= 112.778 \pm 2.89823(3.400) \\ &= (102.92, 122.63) \end{aligned}$$

The same result can be obtained from Minitab by selecting

Stat $\rightarrow$ Basic Statistics $\rightarrow$ 1-Sample t...

Then under "Samples in columns:" select the variable pctidealwgt. Then go to "Options...", set "Confidence level:" to 99, and then hit OK twice. The answer is (102.92, 122.63) which agrees with what we computed by hand.

**Exercises:**

A. Assuming that the population standard deviation for percentage of ideal weight is $\sigma = 15$ in this population, compute a 90% confidence interval for $\mu$.

B. Assuming that the population standard deviation for percentage of ideal weight is unknown, compute a 90% confidence interval for $\mu$.