1. Consider the butterfat production (in pounds) for a cow during a 305-day milk production period following the birth of a calf. Let $X$ and $Y$ equal the butterfat production for such cows on a farm in Wisconsin and a farm in Michigan. Twelve observation of $X$ are  

$$649, 657, 714, 877, 975, 468, 567, 849, 721, 791, 874, 405$$

and sixteen observations of $Y$ are  

$$699, 891, 632, 815, 589, 764, 524, 727, 597, 868, 652, 978, 479, 733, 549, 790$$

(a) Assuming that $X$ is $N(\mu_X, \sigma^2_X)$ and $Y$ is $N(\mu_Y, \sigma^2_Y)$, find a 95% confidence interval for $\mu_X - \mu_Y$.

(b) Find a 90% confidence interval for $\mu_X - \mu_Y$.

(c) Does there seem to be a significant difference in butterfat production for cows on these two farms?

Solution.  

(a) Based on the given data, we have $\bar{x} = 712.25, \bar{y} = 705.4375$, $s^2_x = 29957.8409$, $s^2_y = 20082.1292$ and $s_p = 155.7572$. Note also that $n = 12, m = 16$ and $t_0 = t_{\alpha/2}(n + m - 2) = t_{0.025}(26) = 2.056$. Therefore, the 95% confidence interval for $\mu_X - \mu_Y$ is

$$[712.25 - 705.4375 - (2.056)(155.7572)\sqrt{\frac{1}{12} + \frac{1}{16}}] = [-115.480, 129.105].$$

(b) In this case, $t_0 = t_{\alpha/2}(n + m - 2) = t_{0.05}(26) = 1.706$. So the 90% confidence interval is

$$[712.25 - 705.4375 - (1.706)(155.7572)\sqrt{\frac{1}{12} + \frac{1}{16}}] = [-94.659, 108.284].$$

(c) No, since 0 is included in the confidence intervals.

2. Assume that the birth weight in grams of a baby born in the United States is $N(3315, 525^2)$, boys and girls combined. Let $X$ equal the weight of a baby girl who is born at home in Ottawa County and assume that the distribution of $X$ is $N(\mu_X, \sigma^2_X)$.

(a) Using 11 observations of $X$, give the test statistics and critical region for testing
$H_0: \mu_X = 3315$ against $H_1: \mu_X > 3315$ if $\alpha = 0.01$. 

(b) Calculate the value of the test statistic and give your conclusion using the following weights:

3119 2657 3459 3629 3345 3629 3515 3856 3629 3345 3062

(c) Give the test statistic and critical region for testing $H_0: \sigma^2_X = 525^2$ against $H_1: \sigma^2_X < 525^2$ if $\alpha = 0.05$. Calculate the value of your test statistic and state your conclusion.

(d) Find the limits of $p$-value for the second test.

**Solution.** (a) The test statistic is

$$T = \frac{\bar{X} - 3315}{S/\sqrt{n}}.$$ 

The critical region is $t \geq t_{0.01}(10)$.

(b) From the given data, we have $\bar{x} = 3385.91$ and $s = 336.32$. Since $t = (3385.91 - 3315)/(336.32/\sqrt{11}) = 0.699 < 2.764$, we do not reject $H_0$.

(c) The test statistic is

$$\chi^2 = \frac{(n - 1)S^2}{\sigma^2_0} = \frac{10S^2}{525^2}.$$ 

The critical region is $\chi^2 \leq \chi^2_{0.95}(10) = 3.94$. The value of the test statistic is $\chi^2 = (10(336.32^2))/(525^2) = 4.104$. Since $\chi^2 = 4.104 > 3.94$, we do not reject $H_0$.

(d) Since $3.94 < 4.104 < 4.865$, we find $0.05 < p - value < 0.10$.

3. A proportion, $p$, that many public opinion polls estimate is the proportion of Americans who would say yes to the question “If something were to happen to the President of the United States, do you think that the Vice President would be qualified to take over as President?” In one such random sample of 1022 adults, 388 said yes.

(a) Based on the given data, find a point estimate of $p$.

(b) Find an approximate 90% confidence interval for $p$.

(c) How large a sample should be taken if we want the maximum error of the estimate of $p$ be equal to 0.02 with 95% confidence?

(d) Suppose that in the past, it was believed that $p = 0.40$. It is now claimed that $p$ has decreased. Define the null and alternative hypotheses, a test statistic and critical region with $\alpha = 0.05$. What is the conclusion of the test?

(e) Find the $p$-value for this test. Can you get the same conclusion as in part (d) if you use $p$-value method?

**Solution.** (a) A point estimate of $p$ is $\hat{p} = 388/1022 = 0.3796$.

(b) For $\alpha = 0.10$, we have $z_{\alpha/2} = 1.645$. So an approximate 90% confidence interval for $p$ is $[0.3796 - 1.645\sqrt{(0.3796)(0.6204)/1022}, 0.3796 + 1.645\sqrt{(0.3796)(0.6204)/1022}] = [0.3546, 0.4046]$.

(c) In this part, we have $\alpha = 0.05$ and $\varepsilon = 0.02$. Since historical data is available, we find that the required sample size is

$$n = \frac{z^2_{\alpha/2}\hat{p}(1 - \hat{p})}{\varepsilon^2} = \frac{1.96^2(0.3796)(1 - 0.3796)}{0.02^2} = 2261.78 \approx 2262.$$
(d) We have $H_0 : p = 0.40$ and $H_1 : p < 0.40$. The test statistic is

$$Z = \frac{Y/n - p_0}{\sqrt{p_0(1 - p_0)/n}} = \frac{Y/1022 - 0.40}{\sqrt{0.40(1 - 0.40)/1022}}.$$  

The critical region is $z \leq -z_{0.05} = -1.645$. The value of the test statistic is

$$z = \frac{388/1022 - 0.40}{\sqrt{0.40(1 - 0.40)/1022}} = -1.331.$$  

Since $z = -1.331 > -1.645$, we do not reject $H_0$. There is no strong evidence to support the claim that $p$ has decreased.

(e) $p$-value$= P(Z \leq z) = P(Z \leq -1.331 = 0.0918$. Since $p$-value$> \alpha = 0.05$, we do not reject $H_0$. So we get the same conclusion as in part (d).

4. Let $x$ and $y$ be the ACT scores in social science and natural science for a student who is applying for admission to a small liberal arts college. A sample of $n = 15$ such students yielded the following data

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$x$</td>
<td>$y$</td>
<td>$x$</td>
<td>$y$</td>
<td>$x$</td>
<td>$y$</td>
</tr>
<tr>
<td>32</td>
<td>28</td>
<td>30</td>
<td>27</td>
<td>26</td>
<td>32</td>
</tr>
<tr>
<td>23</td>
<td>25</td>
<td>17</td>
<td>23</td>
<td>16</td>
<td>22</td>
</tr>
<tr>
<td>23</td>
<td>24</td>
<td>20</td>
<td>30</td>
<td>21</td>
<td>28</td>
</tr>
<tr>
<td>23</td>
<td>32</td>
<td>17</td>
<td>18</td>
<td>24</td>
<td>31</td>
</tr>
<tr>
<td>26</td>
<td>31</td>
<td>18</td>
<td>18</td>
<td>30</td>
<td>26</td>
</tr>
</tbody>
</table>

(a) Find the estimated regression line.

(b) Find point estimate for $\sigma^2$.

**Solution.** (a) From the given data, we find the estimates of $\alpha$ and $\beta$ as follows:

$$\hat{\alpha} = \frac{395}{15} = 26.333,$$

$$\hat{\beta} = \frac{9292 - (346)(395)/15}{8338 - (346^2)/15} = 0.506.$$  

Thus the estimated regression line is given by

$$\hat{y} = 26.333 + (0.506)(x - 23.067) = 0.506x + 14.657.$$  

(b) The estimate of $\sigma^2$ is given by

$$\hat{\sigma}^2 = \frac{1}{15} \sum_{i=1}^{15} (y_i - \hat{y}_i)^2 = 14.126.$$  

3