STAT 305 D Final Exam Practice Problems

1. For $X \sim N(\mu, 1/4)$, find the probability that the random interval, $(X-1, X+1)$ covers μ.

$$
\begin{aligned}
P(\mu \text{ in } (X-1, X+1)) &= P(X-1 < \mu < X+1) \\
&= P(-1 < \mu - X < 1) \\
&= P(-1 < X - \mu < 1) \\
&= P\left(\frac{-1}{\sqrt{1/4}} < \frac{X-\mu}{\sqrt{1/4}} < \frac{1}{\sqrt{1/4}}\right) \\
&= P(-2 < Z < 2) \\
&= P(Z \leq 2) - P(Z \leq -2) \\
&= 0.95454
\end{aligned}
$$

2. Vardeman and Jobe chapter 6 exercise 1. The data are available at http://www.will-landau.com/stat305/data/jmp/towels.jmp. Ignore the bit on tolerance intervals.

Consider the breaking strength data of Table 3.6. Notice that the normal plot of these data given as Figure 3.18 is reasonably linear. It may thus be sensible to suppose that breaking strengths for generic towel of this type (as measured by the students) are adequately modeled as normal. Under this assumption,

(a) Make and interpret 95% two-sided and one-sided confidence intervals for the mean breaking strength of generic towels (make a one-sided interval of the form $(\#,\infty)$).

(b) Make and interpret 95% two-sided and one-sided prediction intervals for a single additional generic towel breaking strength (for the one-sided interval, give the lower prediction bound).

(c) Make and interpret 95% two-sided and one-sided tolerance intervals for 99% of generic towel breaking strengths (for the one-sided interval, give the lower tolerance bound).

(d) Make and interpret 95% two-sided and one-sided confidence intervals for σ, the standard deviation of generic towel breaking strengths.

(e) Put yourself in the position of a quality control inspector, concerned that the mean breaking strength not fall under 9,500 g. Assess the strength of the evidence in the data that the mean generic towel strength is in fact below the 9,500 g target. (Show the whole five-step significance-testing format.)

(f) Now put yourself in the place of a quality control inspector concerned that the breaking strength be reasonably consistent—i.e., that σ be small. Suppose in fact it is desirable that σ be no more than 400 g. Use the significance-testing format and assess the strength of the evidence given in the data that in fact σ exceeds the target standard deviation.

(a) The two-sided 95% confidence interval is given by equation (6-20). The required t is $Q(.975)$ of the t_9 distribution, since (by symmetry) there must be probability .025 in each tail. From Table B-4, $t = Q_9(.975) = 2.262$. From the data, $n = 10$, $\bar{x} = 9082.2$, and $s = 841.87$, so the confidence interval is

$$9082.2 \pm 2.262\left(\frac{841.87}{\sqrt{10}}\right) = 9082.2 \pm 602.19 = [8480.0, 9684.4] \text{ g.}$$

To make the 95% one-sided confidence interval, construct a 90% two-sided confidence interval and use the lower endpoint. The appropriate t for a 90% two-sided confidence interval is $t = Q_9(.95) = 1.833$, and so the 95% one sided interval is

$$9082.2 - 1.833\left(\frac{841.87}{\sqrt{10}}\right) = 9082.2 - 487.99 = 8594.2 \text{ g.}$$

(b) Use equation (6-78). The two-sided 95% prediction interval is

$$9082.2 \pm 2.262(841.87)\sqrt{1+\frac{1}{10}} = 9082.2 \pm 1997.25 = [7084.9, 11079.5] \text{ g.}$$

To make the 95% one-sided prediction interval, construct a 90% two-sided prediction interval and use the lower endpoint.

$$9082.2 - 1.833(841.87)\sqrt{1+\frac{1}{10}} = 9082.2 - 1618.5 = 7463.7 \text{ g.}$$

(c) Use formulas (6-83) and (6-85). $p = .99$, and from Table B-7-A, for 95% confidence, with $n = 10$, $\tau_2 = 4.437$. The resulting two-sided tolerance interval is

$$\begin{aligned} 9082.2 \pm 4.437(841.87) &= 9082.2 \pm 3735.37 \\ &= [5346.8, 12817.6] \text{ g.} \end{aligned}$$

From Table B-7-B, for 95% confidence, with $n = 10$ and $p = .99$, $\tau_1 = 3.981$. The resulting one-sided tolerance interval is

$$\begin{aligned} 9082.2 - 3.981(841.87) &= 9082.2 - 3351.48 \\ &= 5730.7 \text{ g.} \end{aligned}$$

(d) Use equation (6-42) and Table B-5. For a 95% two-sided interval, $U = Q_9(.975) = 19.023$ and $L = Q_9(.025) = 2.700$. The resulting interval for σ^2 is $[335314, 2362476]$; taking the square root of each endpoint, the interval for σ is $[579.1, 1537.0]$ g.

For a 95% one-sided interval, $U = Q_9(.95) = 16.919$ and the interval for σ^2 is $[377013, \infty]$; taking the square root, the interval for σ is $[614.0, \infty]$ g.

(e) 1. H_0: $\mu = 9,500$ g.
2. H_a: $\mu < 9,500$ g.

3. The test statistic is

$$T = \frac{\bar{x} - 9,500}{\frac{s}{\sqrt{10}}}$$

and the reference distribution is the t_9 distribution. Observed values of T far below zero will be considered as evidence against H_0.
4. The sample gives

$$t = -1.57$$

5. The observed level of significance is

$$P(\text{a } t_9 \text{ random variable} < -1.57) = P(\text{a } t_9 \text{ random variable} > 1.57)$$

which is between .05 and .1, according to Table B-4. This is moderate evidence that the mean breaking strength of generic towels is less than 9,500 g.

(f) 1. H_0: $\sigma = 400$ g.
2. H_a: $\sigma > 400$ g.
3. The test statistic is

$$X^2 = \frac{(n-1)s^2}{(400)^2}$$

and the reference distribution is the χ^2_9 distribution. Large observed values of X^2 will be considered as evidence against H_0.
4. The sample gives

$$x^2 = 39.87$$

5. The observed level of significance is

$$P(\text{a } \chi^2_9 \text{ random variable} > 39.87)$$

which is less than .005, according to Table B-5. This is very strong evidence that the standard deviation of breaking strengths of generic towels is greater than 400 g.

3. Vardeman and Jobe chapter 6 exercise 2.

The study mentioned in Exercise 5 also included measurement of the outside diameters of the 16 bushings. Two of the students measured each of the bushings, with the results given here.

Bushing	**1**	**2**	**3**	**4**
Student A	.3690	.3690	.3690	.3700
Student B	.3690	.3695	.3695	.3695
Bushing	**5**	**6**	**7**	**8**
Student A	.3695	.3700	.3695	.3690
Student B	.3695	.3700	.3700	.3690

Bushing	**9**	**10**	**11**	**12**
Student A	.3690	.3695	.3690	.3690
Student B	.3700	.3690	.3695	.3695
Bushing	**13**	**14**	**15**	**16**
Student A	.3695	.3700	.3690	.3690
Student B	.3690	.3695	.3690	.3690

(a) If you want to compare the two students' average measurements, the methods of formulas (6.35), (6.36), and (6.38) are not appropriate. Why?

(b) Make a 95% two-sided confidence interval for the mean difference in outside diameter measurements for the two students.

(a) The formulas are for comparing two means based on two independent samples. Because each bushing was measured twice by each student, there is one paired sample here, not two independent samples.

(b) Compute the differences between students A and B for each bushing, and use equation (6-25). (I took the differences as Student A−Student B.) For 95% confidence, the appropriate t is $t = Q_{15}(.975) = 2.131$, from Table B-4.

$$
\begin{aligned}
-00009375 \pm 2.131 \left(\frac{.0004552929}{\sqrt{16}} \right) &= -00009375 \pm .0002414191 \\
&= [-0.0003352, .0001477].
\end{aligned}
$$

Since zero is in this interval, there is no evidence of a mean difference between students.

4. Vardeman and Jobe chapter 6 exercise 15. You will not have to make any normal probability plots on the test.

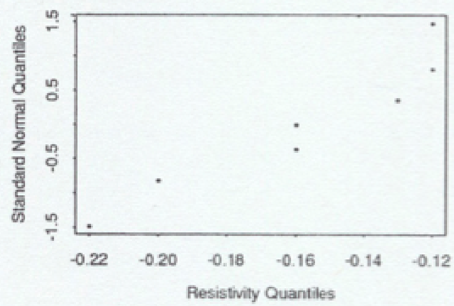
Hamilton, Seavey, and Stucker measured resistances, diameters, and lengths for seven copper wires at two different temperatures and used these to compute experimental resistivities for copper at these two temperatures. Their data follow. The units are $10^{-8}\ \Omega\text{m}$.

Wire	0.0°C	21.8°C
1	1.52	1.72
2	1.44	1.56
3	1.52	1.68
4	1.52	1.64
5	1.56	1.69
6	1.49	1.71
7	1.56	1.72

(a) Suppose that primary interest here centers on the difference between resistivities at the two different temperatures. Make a normal plot of the seven observed differences. Does it appear that a normal distribution description of the observed difference in resistivities at these two temperatures is plausible?

(b) Give a 90% two-sided confidence interval for the mean difference in resistivity measurements for copper wire of this type at 21.8°C and 0.0°C.

(a) I computed the differences as 0.0°C − 21.8°C.



Given the amount of data, there is no evidence that the normal assumption is unreasonable. The plot is roughly linear with no outliers.

222 Chapter

(b) Use equation (6-25). For 90% confidence, the appropriate t is $t = Q_6(.95) = 1.943$, from Table B-4.

$$
\begin{aligned}
-.15857 \pm 1.943 \left(\frac{.03933979}{\sqrt{7}}\right) &= -.15857 \pm .02889055 \\
&= [-.1875, -.1297]\Omega\text{m}
\end{aligned}
$$

5. Vardeman and Jobe chapter 9 exercise 3. The data are available at http://www.will-landau.com/stat305/data/jmp/wirebonding.jmp.

The article "How to Optimize and Control the Wire Bonding Process: Part II" by Scheaffer and Levine (*Solid State Technology*, 1991) discusses the use of a $k = 4$ factor central composite design in the improvement of the operation of the K&S 1484XQ bonder. The effects of the variables Force, Ultrasonic Power, Temperature, and Time on the final ball bond shear strength were studied. The accompanying table gives data like those collected by the authors. (The original data were not given in the paper, but enough information was given to produce these simulated values that have structure like the original data.)

Force, x_1 (gm)	Power, x_2 (mw)	Temp., x_3 °C	Time, x_4 (ms)	Strength, y (gm)

(a) Fit both the full quadratic response surface and the simpler linear response surface to these data. On the basis of simple examination of the R^2 values, does it appear that the quadratic surface is enough better as a data summary to make it worthwhile to suffer the increased complexity that it brings with it? How do the s_{SF} values for the two fitted models compare to s_P computed from the final six data points listed here?

(c) In the linear model $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \epsilon$, give a 90% confidence interval for β_2. Interpret this interval in the context of the original engineering problem. (What is β_2 supposed to measure?) Would you expect the p-value from a test of $H_0: \beta_2 = 0$ to be large or to be small?

(a) The following (abbreviated) printout was made using Minitab Version 9.1.

```
MTB > brief = 3
MTB > regress c1 14 c2-c15

The regression equation is
y = - 1 - 2.30 x1 - 0.08 x2 + 0.836 x3 - 3.99 x4 + 0.0152 x1sq
        + 0.00130 x2sq - 0.00011 x3sq - 0.0078 x4sq + 0.0240 x1*x2
        - 0.0093 x1*x3 + 0.0755 x1*x4 - 0.00467 x2*x3
        + 0.0237 x2*c4 + 0.0007 x3*x4

s = 5.335        R-sq = 81.6%     R-sq(adj) = 64.5%


Analysis of Variance

SOURCE       DF        SS          MS         F        p
Regression   14     1898.33      135.59      4.76    0.002
Error        15      426.93       28.46
Total        29     2325.26

MTB > regress c1 4 c2-c5

The regression equation is
y = - 37.5 + 0.212 x1 + 0.498 x2 + 0.130 x3 + 0.258 x4

Predictor       Coef        Stdev     t-ratio        p
Constant      -37.48        13.10       -2.86    0.008
x1            0.2117       0.2106        1.01    0.324
x2           0.49833      0.07019        7.10    0.000
x3           0.12967      0.04211        3.08    0.005
x4            0.2583       0.2106        1.23    0.231

s = 5.158        R-sq = 71.4%     R-sq(adj) = 66.8%

Analysis of Variance

SOURCE       DF        SS          MS         F        p
Regression    4     1660.14      415.04     15.60    0.000
Error        25      665.12       26.60
Total        29     2325.26

 Obs.      x1         y       Fit   Stdev.Fit   Residual   St.Resid
                               .
                               .
                               .
  24      35.0     41.800    40.990     2.307      0.810       0.18
```

The increase in R^2 going from the simpler to more complex model seems to be small compared to the increase in complexity. Based on the center points, $s_p = 6.016$, so the s_{SF} values for both equations are smaller than s_p. This indicates no problems with either model, but the linear response surface model is simpler.

(c) Use equation (9-47) and the Minitab printout. The appropriate t is $t = Q_{25}(.95) = 1.708$ from Table B-4. The interval for β_2 is

$$\begin{aligned} & .49833 \pm 1.708(.07019) \\ = & .49833 \pm .1198845 \\ = & [.378, .618] \text{ gm.} \end{aligned}$$

Assuming the model is accurate, β_2 represents the mean increase in the final ball bond shear strength that accompanies a 1 mw increase in Power, holding all of the other factors fixed. Since this interval does not contain zero, the p-value for this test would be small.