1) a) We want to find the value $c$ such that

$P(X > c) = .05$ where $X \sim Exp(.5)$

So we want

$$P(X > c) = 0.05$$

$$1 - P(X \le c) = .05$$

$$1 - F(c) = .95$$

$$1 - e^{-c/.5} = .95$$

$$-e^{-2c} = .05$$

$$e^{-c/.5} = .05$$

$$-c/.5 = \log(.05)$$

$$c = -.5 \log(.05)$$

$$c = 1.497$$

b) The type II error rate is the probability we fail to reject the null conditioned on the alternative being true. For us to reject $H_0$ we would need the printer to print before 1.49 minutes have passed. Under the alternative it is impossible to print before 1.49 minutes so we will never make a type II error. This implies

Type II error rate $= 0$

2) a) $\bar{X} = 3.866$

$S = 3.461$

$n = 15$

for a 95% CI we want to use $t_{14,.975} = 2.145$

So our 95% CI is

$$\bar{X} \pm t_{14,.975} \frac{S}{\sqrt{n}}$$

$$\Rightarrow 3.866 \pm (2.145)\left(\frac{3.461}{\sqrt{15}}\right)$$

the interval is
$(1.949, 5.783)$

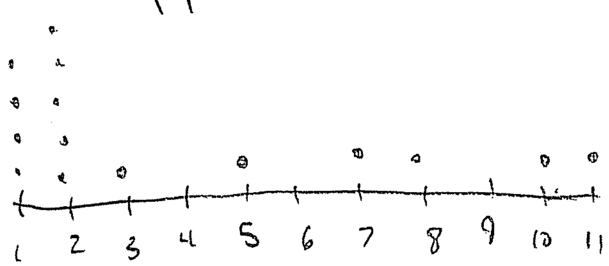
b) $H_0: \mu = 5$
$H_a: \mu \neq 5$

we could do the full 5-step procedure and on the exam I'll be clear about what exactly I want. Here though I will only present the decision/conclusion

Fail to reject $H_0$ since 5 is in our 95% confidence interval.

we do not have evidence that the true mean for this population is different from 5.

c)




This is quite skewed so I don't believe a normal distribution is appropriate here.

3) a)

i) $H_0: \mu_B = \mu_D$
$H_a: \mu_B \neq \mu_D$

ii) we will use $\alpha = .05$

iii) our test statistic will be

$$T = \frac{(\bar{X}_B - \bar{X}_D) - 0}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

where $T \sim T_{n_1+n_2-2}$ assuming $H_0$ is true and the data in both groups are independent and come from normal distributions with a common variance.

Reject $H_0$ if $|T| > t_{.975,4} = 2.776$

iv) $\bar{X}_B = .973$ $\quad S_B = .011547$ $\quad S_p^2 = \frac{(3-1)(.011547)^2 + (3-1)(.02)^2}{3+3-2}$

$\bar{X}_D = 1.04$ $\quad S_D = 0.02$

$S_p^2 = .00026$

$S_p = .01632$

$$T = \frac{(.973 - 1.04)}{.01632 \sqrt{\frac{1}{3} + \frac{1}{3}}} = -5$$

v) with $|T| = 5 > 2.776$ we reject $H_0$ in favor of $H_a$

vi) There is enough evidence to conclude that the two brands have different true mean stretch values.

3 b) Our 95% CI for the difference of the means has the form

$$(\overline{X}_B - \overline{X}_D) \pm t_{.975,4} \; S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$\Rightarrow (.973 - 1.04) \pm (2.776)(.01632)\sqrt{\frac{1}{3} + \frac{1}{3}}$$

$$\Rightarrow -0.0666 \pm .0369$$

$\Rightarrow (-0.1036, -0.029)$ is our 95% CI

c) Yes - they both tell us we have evidence of a difference in the true average stretch values for these two brands.

4) a) $$b_1 = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2} = \frac{2.40}{1.84} = 1.304348$$

$$b_0 = \bar{y} - b_1\bar{x} = \frac{30.53}{20} - (1.304348)(.5) = 0.874326$$

So the least squares line is

$$\hat{y} = 0.874326 + 1.304348x$$

b) $$\hat{y} = 0.874326 + 1.304348(.2) = 1.135$$

$$\hat{y} = 0.874326 + 1.304348(1.3) = 2.569978$$

c) I trust the prediction for $x = .2$ more since 1.3 isn't in the range of the data so the prediction for $x = 1.3$ is extrapolated but the prediction for $x = .2$ isn't extrapolated.

4) d) No - there do not seem to be any noticeable problems.

e) i) $H_0: \beta_1 = 1$

$H_a: \beta_1 \neq 1$

ii) I will use $\alpha = .05$

iii) I will use the test statistic

$$K = \frac{b_1 - 1}{SD(b_1)} = \frac{b_1 - 1}{S_{LF} / \sqrt{\Sigma(x_i - \bar{x})^2}}$$

I assume

- $H_0$ is true
- The model $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$ with $\epsilon_i \overset{iid}{\sim} N(0, \sigma^2)$ is correct.

Under these assumptions $K \sim t_{n-2} = t_{18}$

So we will reject $H_0$ if $|K| > t_{18, .975} = 2.101$

iv) $$K = \frac{1.304348 - 1}{\left(\frac{0.8995}{\sqrt{1.84}}\right)} = 0.4589$$

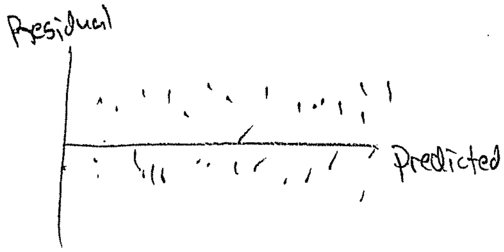
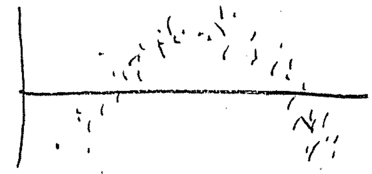
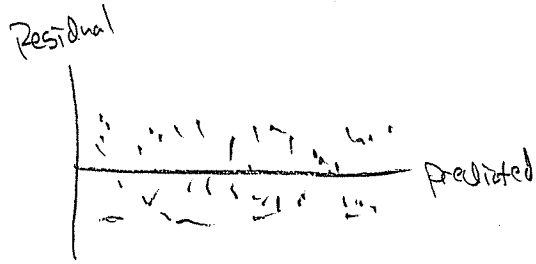
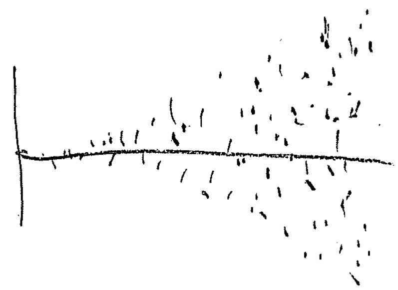
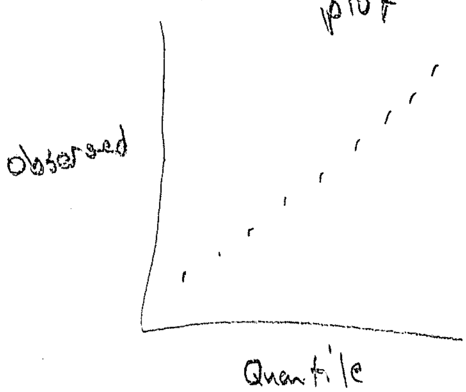
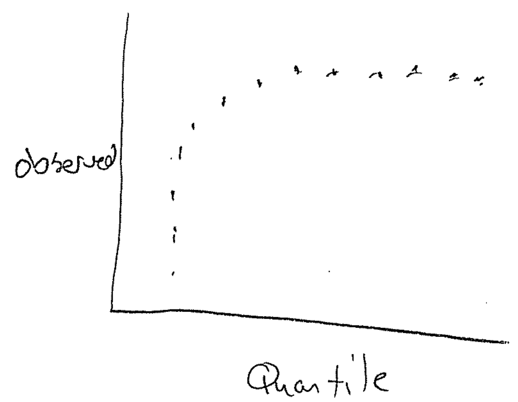
v) with $|K| = 0.4589 < 2.101$ we fail to reject $H_0$

vi) We do not have evidence (at the .05 level) that the true slope is different from 1.

5)

| Assumption | Diagnostic |
| --- | --- |
| 1) Correct model | Check the residual by predicted plot. If the assumption is met there shouldn't be any left over pattern. |
| 2) Constant variance | Check the residual by predicted plot. we want the spread to be constant for all predicted values. If there is a "fan shape" then we might believe this assumption is broken. |
| 3) Normally distributed error terms | we use a normal quantile plot of the residuals (or standardized residuals). If the assumption is met then the points should fall on a straight line. |
| 4) Independence | This is the hardest to check since there isn't an easy diagnostic for this assumption. we need to rely on our knowledge about the data to assess this assumption. |

6)

| Assumption | Good plot | Bad plot |
|---|---|---|
| Correct Model |  |  |
| Constant variance |  |  |
| Normality of error term | Normal quantile plot  |  |

7 a) We know the t-ratio is $\frac{\text{Estimate}}{\text{Std Error}} = \frac{.1050245}{0.007365} = 14.25995$

we have 242 degrees of freedom and the largest quantile the book gives for a $t$ w/ 242 df is

$Q(.9995) = 3.373$

Since $14.25995 > 3.373$ we know

$$P(T_{242} > 14.25995) < P(T_{242} > 3.373) = .0005$$

so $P(|T_{242}| > 14.25995) = 2 \cdot P(T_{242} > 14.25995) < 2(.0005) = .0001$

So our p-value here is $< .0001$.

b) $H_0: B_1 = 0.10$

$H_a: B_1 \neq 0.10$

$$T = \frac{b_1 - 0.10}{SE(b_1)} = \frac{0.1050245 - 0.10}{0.007365} = 0.6822132.$$

It doesn't matter too much what we choose for $\alpha$ in this case we are going to fail to reject $H_0$

- we don't have evidence that for every increase of a dollar in total bill that the increase in tip is different than 10 cents.

c) No – 0 won't be in the interval since the test $H_0: B_0 = 0$ $H_a: B_0 \neq 0$ was significant at the 0.05 level.

d) $SE(\text{intercept}) = \frac{\text{Estimate}}{t\text{-ratio}} = \frac{.9202696}{5.76} = .159769$ (since $t\text{-ratio} = \frac{\text{Est}}{\text{Std Err}}$)

our 95% CI has the form

$$\text{intercept} \pm t_{.975,\,242}\, SE(\text{intercept})$$

$$.9202696 \pm (1.98)(.159769)$$

$$\Rightarrow (.6039269,\ 1.2366123)$$

e) 0 isn't in the range of the x-values so this is extrapolating. I'm not sure it makes sense for a tip between 60 cents and \$1.23 to be given with a bill of \$0 either. I would be wary of taking the confidence interval too seriously.

f) Our prediction $\mu_{y|x=25} = 0.9202696 + 0.1050245(25) = 3.545882$.

the standard error of this prediction is: $S_{LF}\sqrt{\frac{1}{n}\frac{(x-\bar{x})^2}{\sum(x_i-\bar{x})^2}}$

$$= \underbrace{\left(\frac{S_{LF}}{\sqrt{\sum(x_i-\bar{x})^2}}\right)}_{SE(b_1)}\sqrt{\frac{(x-\bar{x})^2}{n}}$$

so we just need $\bar{x} = \frac{\bar{y}-b_0}{b_1} = \frac{2.998279 - 0.9202696}{0.1050245} = 19.78595$

$b_0 = \bar{y} - b_1\bar{x}$

$$= (0.007365)\sqrt{\frac{(25-19.78595)^2}{244}} = .002458403$$

our CI has the form $\hat{y} \pm t_{242,\,.975}\, SE(\hat{y})$

$$\Rightarrow 3.545882 \pm (1.98)(.002458403) \Rightarrow (3.54101,\ 3.55075)$$

... Although I did most of that because I forgot that I gave $\Sigma(x_i - \bar{x})^2 = 19258.46$

So if we use that then we see the standard error is: $S_{LF}\sqrt{\frac{1}{n}\frac{(x-\bar{x})^2}{\Sigma(x_i-\bar{x})^2}} = 1.022048\sqrt{\frac{1}{244}\frac{(25-19.78545)^2}{19258.46}}$

$= 0.002459833$

which is the same (essentially)