STA 108 HW 2 - Ishita Dutta

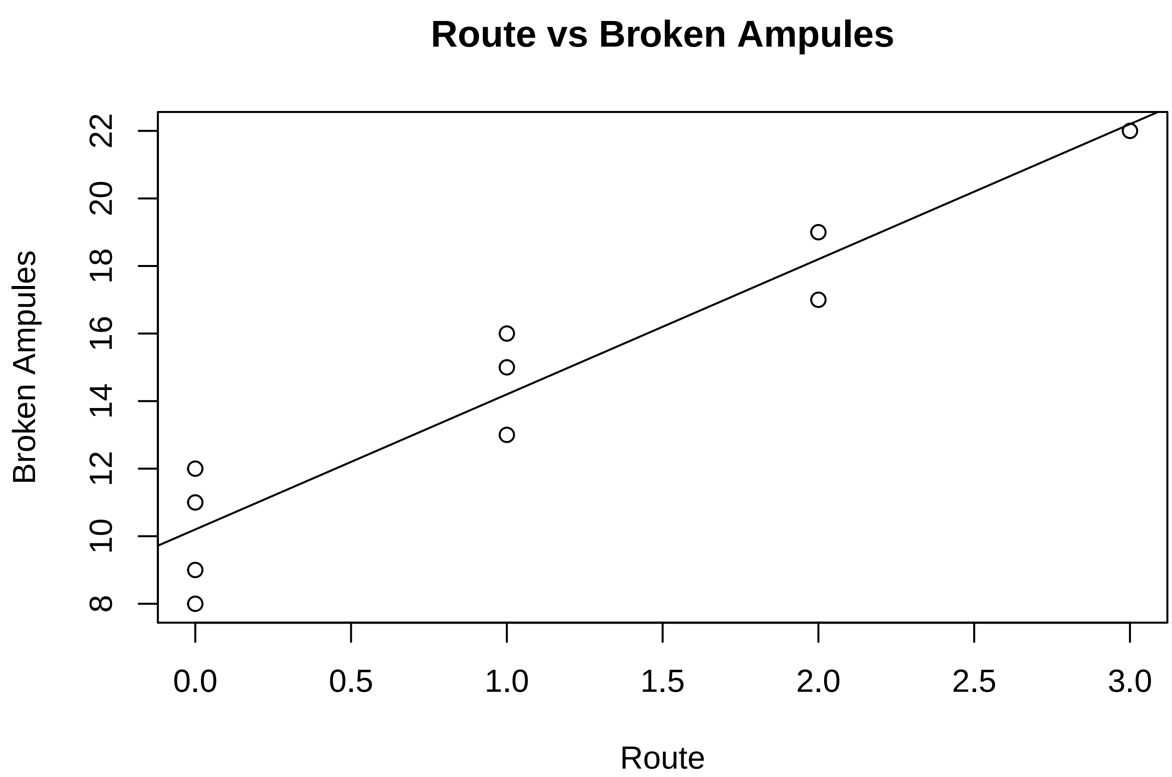
1.21

a)

```r
route <- c(1, 0, 2, 0, 3, 1, 0, 1, 2, 0)
broken <- c(16, 9, 17, 12, 22, 13, 8, 15, 19, 11)
b1hat = t(route-mean(route))%*%(broken-mean(broken))/sum((route-mean(route))^2)
b0hat = mean(broken) - b1hat*mean(route)
cat("Y = ", b0hat, " + ", b1hat, "x")
```

```
## Y =  10.2  +  4 x
```

```r
plot(route, broken, xlab="Route", ylab="Broken Ampules" , main = "Route vs Broken Ampules")
abline(lm(broken ~ route))
```



The regression appears to be a good fit here, as the line has an even amount of points to either side of it and is equally close to the points on either side b)

```r
cat("Point estimate at X = 1: ", b0hat + (b1hat * 1), " broken ampules.")
```

```
## Point estimate at X = 1:  14.2  broken ampules.
```

c) An estimate would be the same as B1, so 4.

d)

```r
x_bar = mean(route)
y_bar = mean(broken)
y_bar
```

```
## [1] 14.2
```

```r
Point = b0hat + b1hat*(x_bar)
Point
```

```
##      [,1]
## [1,] 14.2
```

1.29

This would mean that the intercept of the graph is at the origin, and for the regression line, it means that the entire line is only counting growth and error. Starting point is not counted here, since it is 0.

2.1

a) The conclusion is warranted, since the confidence interval for the slope is all above 0. Since this is at a significance of 0.05, the slope is significantly different from 0.

b) The intercept X = 0 holds no significance since that is not in the scope of the situation.

2.3

The student's statement is inaccurate as we are testing for the slope to equal to zero or be different from zero. The p-value of 0.91 means that we fail to reject the null hypothesis and cannot conclude that such a linear relationship exists in this context.

2.7

```r
cat("a)\n")
```

```
## a)
```

```r
x = c(16, 16, 16, 40, 40, 40)
y = c(199, 205, 196, 248, 253, 246)
b1hatval = t(x-mean(x))%*%(y-mean(y))/sum((x-mean(x))^2)
b0hatval = mean(y) - b1hatval*mean(x)
y_mean = b0hatval + b1hatval*(2)
variance = ((1/(5)) * (sum ((y - mean(y)) ^ 2))) / ((1/(5)) * (sum ((x - mean(x) ) ^ 2)))
meval = (qnorm(0.99))*(sqrt(variance)/sqrt(5))
cat("99% Confidence Interval for Mean at X = 2: [", y_mean - meval, ", ", y_mean + meval, "] \nb)\n")
```

```
## 99% Confidence Interval for Mean at X = 2: [ 169.2726 ,  173.5607 ]
## b)
```

```r
interval = var.test(x, y, alternative = "two.sided", conf.level = 0.99)
cat("Conclusion: cannot conclude that there is a significant difference in variances.\nc)\n")
```

```
## Conclusion: cannot conclude that there is a significant difference in variances.
## c)
```

```r
delta = 0.3 / 0.1
power = pt(q = meval, df = 1, ncp = delta, lower.tail = FALSE) + pt(q = meval, df = 6 - 2, ncp = delta)
power
```

```
## [1] 1.007328
```

2.15 (a, b, c only)

a) The mean breakage is Y for that X, therefore:

```r
mean_2 = b0hat + (b1hat * 2)
mean_4 = b0hat + (b1hat * 4)
cat("Mean at X = 2: ", mean_2, " broken ampules.\n")
```

```
## Mean at X = 2:  18.2  broken ampules.
```

```r
cat("Mean at X = 4: ", mean_4, " broken ampules.\n")
```

```
## Mean at X = 4:  26.2  broken ampules.
```

```r
var = ((1/(10 - 1)) * (sum ((broken - y_bar ) ^ 2))) / ((1/(10 - 1)) * (sum ((route - x_bar ) ^ 2)))
me = (qnorm(0.99))*(sqrt(var)/sqrt(10))
cat("99% Confidence Interval for Mean at X = 2: [", mean_2 - me, ", ", mean_2 + me  , "] \n")
```

```
## 99% Confidence Interval for Mean at X = 2: [ 15.09975 ,  21.30025 ]
```

```r
cat("99% Confidence Interval for Mean at X = 4: [", mean_4 - me, ", ", mean_4 + me  , "] ")
```

```
## 99% Confidence Interval for Mean at X = 4: [ 23.09975 ,  29.30025 ]
```

With the confidence interval, we are 99% confident that the number of ampules broken at 2 and 4 transfers are within those intervals, respecively.

b) The confidence interval is the same as part (a): [15.09975 , 21.30025]. This is the range with 99% confidence of the number of ampules that will break at the two transfers in that shipment.

c) Because the shipments are independent, the confidence interval for each of them will be [15.09975, 21.30025]. To get the total for 3 shipments, we add the bounds: [45.29925, 63.90075]