I apologize that it is not as well put together this week. I did not have time to make it look better. Thanks!

Problem 1

a) The variable sex is significant at a 0.05 significance level. The variable income is significant at a 0.01 significance level.
b) As the variable "sex" is increased by one unit, it is estimated that the gamble variable will change by -22.11833 units (as long as all other variables are held constant). Because the variable only takes values of 0 and 1, this means that it is estimated to drop -22.11833 as it switches from sex = 0 to sex = 1 (as long as all other variables are held constant).
c) F = 4.1338, p-value = 0.01177. Statistically significant at significance level 0.05.

Problem 2

a) $H_0: \beta_{salary} = 0; \;\; H_A: \beta_{salary} \neq 0$
   T = -1.878
   Pval = 0.0667
   Statistically significant at the 0.10 significance level

   $H_0: \beta_{salary} = \beta_{ratio} = \beta_{expend} = 0; \;\; H_A: at\ least\ one\ is\ not\ 0$
   F = 4.0662
   Pval = 0.01209
   Statistically significant at the 0.05 significance level

b) $H_0: \beta_{salary} = 0; \;\; H_A: \beta_{salary} \neq 0$
   T = 0.686
   Pval = 0.496
   Fail to reject null. We do not have statistically significant evidence that $\beta_{salary} \neq 0$

   F = 52.88
   Pval is approximately 0.
   Reject the null, statistically significant at and reasonable significance level.

   The F-Test and t-test are not equivalent. Different decisions are made between rejecting the null and failing to reject the null at different significance levels.

Problem 3

Lcavol only: Residual standard error = 0.7875 and R-squared = 0.5394

Lweight added: Residual standard error = 0.7506 and R-squared = 0.5859

SVI added: Residual standard error = 0.7168 and R-squared = 0.6264

Lpbh added: Residual standard error = 0.7108 and R-squared = 0.6366

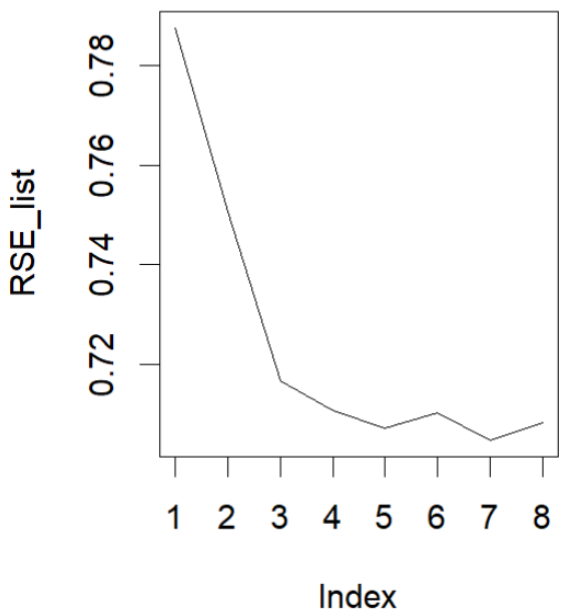
Age added: Residual standard error = 0.7073 and R-squared = 0.6441

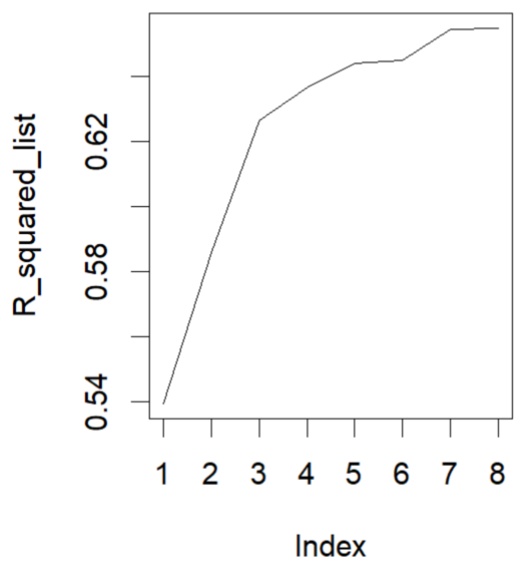
LCP added: Residual standard error = 0.7102 and R-squared = 0.6451

PGG45 added: Residual standard error = 0.7048 and R-squared = 0.6544

Gleason added: Residual standard error = 0.7084 and R-squared = 0.6548

As the number of variables increased the RSE decreased on average, but R^2 increased. Generally, as one of the quantities went up, the other went down.





Problem 4

$y_{ij} \sim N(\mu + \alpha_i, \sigma^2)$

$\mu = 0$ and $\alpha_1 = \frac{y_{1\cdot}}{5}$

$Var(\hat{\mu} + \widehat{\alpha_1}) = \frac{\sigma^2}{5}$

$\hat{\mu} + \widehat{\alpha_1} \sim N(\mu + \alpha_1, \frac{\sigma^2}{5})$

R Code:

```
library(faraway)

#Problem 1

out_full <- lm(gamble ~ sex + status + income + verbal, data = teengamb)
summary(out_full)
RSS_full <- sum(out_full$residuals^2)

out_red <- lm(gamble ~ income, data = teengamb)
summary(out_red)
RSS_red <- sum(out_red$residuals^2)

anova(out_full, out_red)

F <- ((RSS_red - RSS_full)/3) / (RSS_full/42)
F

p_value <- pf(F, lower.tail = FALSE, df1 = 3, df2 = 42)
p_value

#Problem 2

out_1 <- lm(total ~ expend + ratio + salary, data = sat)
out_2 <- lm(total ~ 1, data = sat)

anova(out_1, out_2)

out_3 <- lm(total ~ expend + ratio + salary + takers, data = sat)
summary(out_3)

#Problem 3

out_pro_1 <- lm(lpsa ~ lcavol, data = prostate)
summary(out_pro_1)

out_pro_2 <- lm(lpsa ~ lcavol + lweight, data = prostate)
summary(out_pro_2)

out_pro_3 <- lm(lpsa ~ lcavol + lweight + svi, data = prostate)
summary(out_pro_3)

out_pro_4 <- lm(lpsa ~ lcavol + lweight + svi + lbph, data = prostate)
summary(out_pro_4)
```

```
out_pro_5 <- lm(lpsa ~ lcavol + lweight + svi + lbph + age, data = prostate)
summary(out_pro_5)

out_pro_6 <- lm(lpsa ~ lcavol + lweight + svi + lbph + age + lcp, data = prostate)
summary(out_pro_6)

out_pro_7 <- lm(lpsa ~ lcavol + lweight + svi + lbph + age + lcp + pgg45,
        data = prostate)
summary(out_pro_7)

out_pro_8 <- lm(lpsa ~ lcavol + lweight + svi +
        lbph + age + lcp + pgg45 + gleason,
        data = prostate)
summary(out_pro_8)

RSE_list <- c(0.7875, 0.7506, 0.7168, 0.7108, 0.7073, 0.7102, 0.7048, 0.7084)
plot(RSE_list, type = "l")

R_squared_list <- c(0.5394, 0.5859, 0.6264, 0.6366, 0.6441, 0.6451,0.6544, 0.6548)

plot(R_squared_list, type = "l")
```