Math 4397/6397
Problem Set 7, due October 13

Solutions:

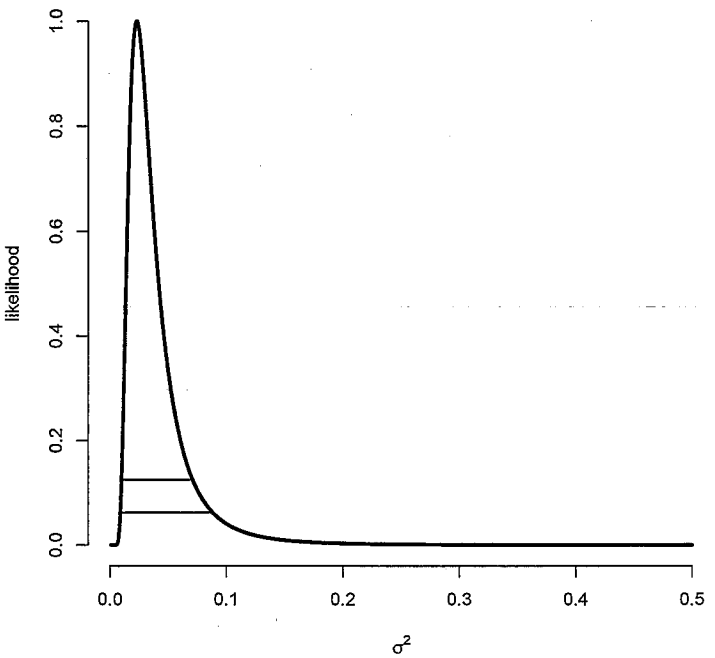
Problem 1. a. Since we are asking about an estimate for the change in FEV, we compute the differences between FEVs of each person and find a confidence interval using the t-distribution:

```
fev1 <- c(3.22, 4.06, 3.85, 3.50, 2.80,
          3.25, 4.20, 3.05, 2.86, 3.50)
fev2 <- c(2.95, 3.75, 4.00, 3.42, 2.77,
          3.20, 3.90, 2.76, 2.75, 3.32)
diff <- fev2 - fev1
round(t.test(diff)$co, 3)
```

The t confidence interval for change in FEV is $[-.255, -.039]$ with a midpoint of $-.147$ suggesting a negative change in FEV volume over the time scale considered.

b. We create a likelihood function for the variance of the change in FEV.

```
s2 <- var(diff)
n <- length(diff)
alpha <- .05
qtiles <- qchisq(c(alpha/2, 1 - alpha/2), n - 1)
ival <- rev((n - 1) * s2 / qtiles)
sqrt(ival)
sigmasqVals <- seq(0.001, .50, length = 1000)
likeVals <- dgamma((n - 1) * s2,
                   shape = (n - 1)/2,
                   scale =  2 * sigmasqVals)
likeVals <- likeVals / max(likeVals)
plot(sigmasqVals, likeVals,
     type = "l",
     frame = F,
     xlab = expression(sigma^2),
     ylab = "likelihood", lwd = 3)
lines(range(sigmasqVals[likeVals >= 1 / 8]), c(1 / 8, 1 / 8))
lines(range(sigmasqVals[likeVals >= 1 / 16]), c(1 / 16, 1 / 16))
range(sigmasqVals[likeVals >= 1 / 16])
```



The likelihood functions is above $1/8$ in the range from is .01 to .07, and above $1/16$ in the range from is .01 to .09.

Problem 2. We compute the independent group t-confidence interval where we compare the differences for smokers to non-smokers, assuming the variances are equal for both groups:

```
fev1smoker <- c(2.85,3.32,3.01,2.95,2.78,2.86,2.78,2.90,
+                2.76,3.00,3.26,2.84,2.50,3.59,3.30)
> fev2smoker <- c(2.88,3.40,3.02,2.84,2.75,3.20,2.96,2.74,
+                  3.02,3.08,3.00,3.40,2.59,3.29,3.32)
> diffsmoker <- fev1smoker - fev2smoker
> diffsmoker <- fev2smoker - fev1smoker
> t.test(diff, diffsmoker, var.equal = TRUE)
```

The mean difference for smokers was .05 while it was $-.147$ for non-smokers. The confidence interval using the pooled estimate for the standard deviaton was $-.37$ to -0.03, which suggests that the smokers do not lose as much FEV as the non-smokers. This confidence interval depends on normality and the assumption of equal variances in the two groups.

Problem 3. Assuming that growth in each month is independent, identically distributed for plants, we have for one month the expected value $\hat{\mu}_A = 9.78/3 = 3.26$ and sample SD $\hat{s}_A = 7.51/3 = 2.50$ and $\hat{\mu}_D = 12.83/3 = 4.28$ and $\hat{s}_D = 8.31/3 = 2.77$. A 95% confidence interval for the average difference in monthly growth increment is then:

$$(3.26 - 4.28) \pm 1.96 \times \sqrt{(2.50^2/260 + 2.77^2/289)} = (-1.46, -0.58)$$

Ignoring biological considerations, the annual difference in growth can then be expected to be 12 times the monthly difference, within the range from -17.5 to -07.0 when comparing fertilizer A to D. So D appears to perform better!