Solutions to HW5, Stat 401 Spring 2011

(1). Extra Problem on web-page (a). This one sided P-value is $1 - \Phi((24 - 22.7)\sqrt{31}/8) = 0.183$.
(b). For power, take $1 - \Phi((22.7 + z_{.05} 8/\sqrt{31} - 25)/(8/\sqrt{31})) = .482$.
(c) Now we find as p-value: $1 - F_{t_{30}}((24 - 22.7)\sqrt{31}/8) = 0.186$. In calculating the power according to the t_{30} distribution, we use `pt(.,30, ncp=1.601)` in **R**, finding power $= 1 - pt(1.697, 30, ncp = 1.601) = 0.468$, where $t_{30,.05} = 1.697$ and the non-centrality parameter was calculated (using $\sigma^2 = 8$) as $\sqrt{31}(25 - 22.7)/8 = 1.601$ (or you could have used the power charts given in the back of the book).

(2). #62, p. 363. Now $n_1 = 48, n_2 = 45, s_1 = 21.5, s_2 = 19.45$. Again we must assume normally distributed samples to justify the use of the F test. The test is now two-sided, and the test statistic of $(21.5/19.45)^2 = 1.222$ corresponds to 2-sided P-value `2*(1-pf(1.222,47,39)) = .504`, so we accept the hypothesis of equality. Note that $F_{47,44,.05} = 1.64$, so we certainly accept at level $\alpha = .10$.

(3). #64, p. 364. Upper 95% confidence bound based on F for two-sample ratio of variances (triacetate on top, versus cotton on bottom), using data from example 9.6, is $(3.59/0.79)^2/F_{9,9,.95} = (3.59/0.79)^2 \cdot qf(.95, 9, 9) = (3.59/0.79)^2(3.179) = 65.65$. That is the (really high!) upper bound on the likely ratio of variances. Since the problem asks for upper bound on ratio of standard deviations, take square root to get 8.10.

(4). #68, pp. 364-5. These two samples, of respective sizes 24, 11, yield $\bar{x} = 103.66$, $s_x = 3.738$, $\bar{y} = 101.11$, $s_y = 3.603$. Here we begin by forming the two-sided 95% CI for ratio of variances:
`(3.738/3.603)`^{2}`/c(qf(.95,23,10),qf(.05,23,10)) = (0.392, 2.448)`.
Therefore (since 1 is close to the middle of the interval) it seems reasonable to do our two-sample t interval using pooled variance, giving $103.66 - 101.11 \pm sqrt((23{*}3.738^2{+}10{*}3.603^2)/33){*}sqrt(1/24{+}1/11){*}2.035 = (-0.189, 5.289)$ as the CI for $\mu_x - \mu_y$. This *does* require normally distributed observations, *iid* within samples.

(5). #6, p. 575. This is a multinomial goodness of fit testing problem with simple null hypothesis $\mathbf{p} = (9/16, , 3/16, 4/16)$ and $n = 368$. Based on the given data, we form the chi-square test statistic $(195-368*9/16)^2/(368*9/16)+(73-368*3/16)^2/(368*3/16)+(100-368*4/16)^2/(368*4/16) = 1.623$. We compare this with a χ^2_2 d.f.: the p-value (probability that such a r.v. exceeds 1.623) is .444, so we do not reject at any reasonable significance level.

(6). #9, pp. 575. Again we have a simple null (only $\lambda = 1$ within H_0).
(a). First we find 5 equal-probability intervals with endpoints $0,\ a_1,\ \ldots, a_4,\ \infty$ by solving the equations

$$k/5 = F(a_k) = \int_0^{a_k} e^{-x} dx = 1 - e^{-a_k} \quad \Rightarrow \quad a_1 = -\log(1 - k/5)$$

which implies that the a_k values are (0.223, 0.511, 0.916, 1.609).
(b). Next, tally the given observations, counting how many fall in the 5 class intervals $[0, .223], (.223, .511], \ldots, (1.609, \infty)$, giving $(N_1, N_2, \ldots, N_5) = (6, 8, 10, 7, 8)$. The corresponding expected count in each interval is $0.2*40 = 8$, so the test statistic value is $((6-8)^2 + (8-8)^2 + (10-8)^2 + (7-8)^2 + (8-8)^2)/8 = 1.125$, which is much smaller than $\chi^2_4 = 9.48$.
So we accept H_0.

(7). #72, p. 365. Here we take the $n = 17$ *differences* D_i as our basic data in a paired analysis of mean difference, obtaining $\bar{D}) = -4.176$, $s_D = 35.849$. So **no**, the mean difference is not estimated with great precision. Relying on the qqplot (not required that you do it yourself for this problem) we give the mean difference as a (really wide !) 95% confidence interval $-4.176 \pm 35.489 * 2.12/sqrt(17) = (-14.07,\ 22.42)$.

(8). #94, p. 179. Here we do three normal probability plots, first for original data, then square roots, then cube roots:

```
> Ex4.94 = scan("ex04-94.txt", skip=1)
> qqnorm(Ex4.94) ### not too good, vaguely quadratic rather than linear
> qqnorm(sqrt(Ex4.94)) ### nearly perfect, except for points at end
> qqnorm(Ex4.94^(1/3)) ### this one is even better
```