HW 10

7.57 **(a)** We test H_0: $\mu_1 = \mu_2$ vs. H_a: $\mu_1 < \mu_2$. $t = -7.34$, which gives $P < 0.0001$ whether df = 133 or df = 140.6. Cocaine use is associated with lower birth weights. **(b)** The standard error of the difference is $SE_D \doteq 52.47$, and the interval is $(\bar{x}_1 - \bar{x}_2) \pm t^* SE_D$. Answers will vary with the degrees of freedom used; see the table. **(c)** The "Other" group may include drug users, since some in it were not tested. Among drug users, there may have been other ("confounding") factors that affected birthweight. Note that in this situation, an experiment is out of the question.

df	t^*	Interval
100	1.984	−489.1 to −280.9 g
133	1.9780	−488.8 to −281.2 g
140.6	1.9770	−488.7 to −281.3 g

8.15 $\hat{p} = \frac{13}{75} = 0.17\overline{3}$, and $SE_{\hat{p}} = \sqrt{\hat{p}(1-\hat{p})/75} = 0.0437$, so the 95% confidence interval is $0.17\overline{3} \pm (1.96)(0.0437)$, or 0.0877 to 0.2590.

8.38 Note that the rules of thumb for the normal approximation are not satisfied here (the number of birth defects is less than 10). Additionally, one might call into question the assumption of independence, since there may have been multiple births to the same set of parents included in these counts (either twins/triplets/etc., or "ordinary" siblings).

If we carry out the analysis in spite of these issues, we find $\hat{p}_1 = \frac{16}{414} \doteq 0.03865$ and $\hat{p}_2 = \frac{3}{228} \doteq 0.01316$. We might then find a 95% confidence interval: $SE_D \doteq 0.01211$, so the interval is $\hat{p}_1 - \hat{p}_2 \pm (1.96)(0.01211) = 0.00175$ to 0.04923. (Note that this does not take into account the presumed direction of the difference.) We could also perform a significance test of H_0: $p_1 = p_2$ vs. H_a: $p_1 > p_2$: $\hat{p} = \frac{19}{642} \doteq 0.02960$, $s_p \doteq 0.01398$, $z \doteq 1.82$, $P = 0.0344$.

Both the interval and the significance test suggest that the two proportions are different, but we must recognize that the issues noted above make this conclusion questionable.

9.12 **(a) & (b)** See table. Percentage of children receiving tetracycline seems to rise as we move from urban to rural counties. **(c)** H_0: There is no relationship between county type and prescription practice; H_a: There is a relationship. **(d)** $X^2 = 7.370 + 0.372 + 7.242 + 5.440 + 0.275 + 5.345 = 26.044$, df = 2, $P < 0.0005$. The differences between the tetracycline prescription practices are significant; doctors in rural counties were most likely to prescribe tetracycline to young children, while urban doctors were least likely to do so.

	Urban	Intermed.	Rural	
Tetra.	65 90.88 30.4%	90 95.98 39.8%	172 140.14 52.1%	327 42.5%
No tetra.	149 123.12 69.6%	136 130.02 60.2%	158 189.86 47.9%	443 57.5%
	214	226	330	770