Math 159, Spring 2005
Professor Hardin
Sample Exam #1
Solutions

1. (a) We'll consider height to be continuous. The question of interest is testing whether the 20^{th} quantile is more than 100 lbs. We can translate this question into:
$H_o : p = P(Y \leq 100) \geq 0.2$
$H_1 : p < 0.2$
Our test statistic is $T_1 = T_2 = 4, n = 16$. p-value $= P(Y \leq 4|p = .2) = 0.7982$. We can't reject the null hypothesis. (To keep your test within $\alpha = 0.05$, your critical region should be – CR: $T \leq 0$.) There is no evidence that 80% of tenth grade boys weigh more than 100 lbs.

(b) Large sample theory: $\hat{p} \pm z_{1-\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = (0.538, 0.962)$.
Small sample theory: letting $\alpha/2 = \sum \binom{n}{i}(p_1^*)^n(1-p_1^*)^{n-i} \rightarrow (0.476, 0.927)$.
Both confidence intervals give intervals within which we are 95% confident that the true proportion of boys over 100 lbs lies. In general, we use the normal approximation when we see at least 5 successes and 5 failures (and really, we hope for 10!) Here we see only 4 boys with weight under 100 lbs, so we suspect that the normal approximation is less good. The exact values are likely much more accurate.

(c) To test that the median weight is 110 lbs, we write the hypotheses as: $H_o : p = P(Y \leq 110) = 0.5$
$H_1 : p \neq 0.5$
$T_1 = T_2 = 5$. To find the p-value, we need to calculate whether it is more extreme to be bigger than 5 or less than 5 given the null hypothesis:
$P(Y \leq 5|p = 0.5) = 0.9183$
$P(Y \geq 5|p = 0.5) = 0.2018$
p-value $= 2 \cdot P(Y \geq 5|p = 0.5) = 0.4036$
Again, we can't reject our null hypothesis. There is no evidence that the median is different from 110 lbs.

(d) CI for $x_{.25}$:
$\alpha_1 = \sum_{i=0}^{r-1} \binom{16}{i}(.25)^{16}(.25)^{16-i} = 0.0635 = P(Y \leq 1|p = .25)$
$1 - \alpha_2 = \sum_{i=0}^{s-1} \binom{16}{i}(.25)^{16}(.25)^{16-i} = 0.9729 = P(Y \leq 7|p = .25)$
$1 - \alpha = 0.9094 \rightarrow r = 2, s = 8, 90.94\%$ CI for $x_{.25} : (92, 122)$
We are 90.94% confident that the true 25^{th} quantile of tenth-grade boys' heights is between 92 lbs and 122 lbs.

(e) If the diet is an effect means of losing weight, then there will be more students who lose weight than gain or stay the same.
$H_o : P(+) \leq P(-)$
$H_1 : P(+) > P(-)$
p-value $= P(Y \geq 11|n = 15, p = .5) = 0.0107$. We can reject the null hypothesis and claim that this diet is an effective means of losing weight for tenth-grade boys.

2. (a) H_o : the teams are equally good at the game
   H_1 : one team is better at the game
   (b) $P(X \geq 4|p = .5) = \binom{5}{4}(.5)^5 + (.5)^5 = 0.1875$
   (c) You would need to know the extent of the difference (that is, what is the alternative hypothesis really saying?)

3. $P(X = 4) = 0.3$. The 47^{th} quantile is 6. Check: $P(X \leq 6) = .7 \geq .47, P(X < 6) = .45 \leq .47$.

4. (a) Let's say that we're doing a one sided test where we reject for small values of T = total number os $+'s$. That means, let's say, that we reject if $T \leq t$ where: $P(Y \leq t|n, p = .5) = .025$. We can approximate the distribution of Y using the CLT: $Y \sim N(np, np(1-p)) = N(n/2, n/4)$.
   So, to find our critical region, we can calculate the above probability usting:
   $P(Y \leq t) = P(Z \leq \frac{t-np}{\sqrt{np(1-p)}} = P(Z \leq \frac{t-n/2}{\sqrt{n/4}}) = P(Z \leq \frac{2t-n}{\sqrt{n}}) = 0.025$
   Or, $\frac{2t-n}{\sqrt{n}} = z_{.025} \approx -2 \rightarrow t = \frac{n}{2} - \sqrt{n}$
   (b) This approximation holds when we have a large enough sample (at least 5 successes and 5 failures, though 10 of each is better), and when we're interested in $\alpha = 0.05$ for a two sided test and $\alpha = 0.25$ for a one sided test.

5. We let Y = # of observations ≤ 47 in a sample of size 20. Note that we'll reject the null hypothesis (in favor of the alternative) if Y is particularly small.

   If the size of the test is $\alpha = 0.0577$, then we know the critical region must be: $T \leq 6$. So, the power $= 1 - \beta = P(Y \leq 6|p = 0.3 = H_1) = 0.6080$.
   (b) The approximate power is: $P(Y \leq 6|p = .3) = P(Z \leq \frac{6-6}{\sqrt{np(1-p)}} = P(Z \leq 0) = 0.5$.

6. (a) $H_o : P(+) \leq P(-)$
   $H_1 : P(+) > P(-)$
   T = # of +'s = 5, n=6
   We reject if $T \geq t$ where $P(Y \geq t|p = .5) = \alpha$. It turns out that $P(Y \geq 6|p = .5) = 0.0156$ and $P(Y \geq 5|p = .5) = 0.0625$. So, our p-value is $P(Y \geq 5|p = .5) = 0.0625$ which gives some evidence that training can increase hypnotic suseptibility.
   (b) Changing Y_3 to be larger won't have an effect on the test statistic. However, changing any of the other values to switch the sign of the difference would have an effect (it would make it stronger if the - became + and much less strong if any of the +'s became -'s.)

7. (a) 2^{10}
   (b)

1	2	3	4	5	6	7	8	9	10
f	f	m	m	m	m	f	m	f	f

(c) Each of the 2^{10} points in the sample space has an equal probability of happening.

(d) $(.5)^{10} = 0.00098$

(e) $10(.5)^{10} = 0.0098$

(f) $P(X = 3) = \binom{10}{3}(.5)^{10} = 0.1172$

(g) $P(X \leq 3) = 0.00098 + 0.0098 + \binom{10}{2}(.5)^{10} + 0.1172 = 0.172$

8. (a) $\mu = 0.48, \sigma = \sqrt{4(.12)(.88)} = 0.65$

(b) Median is 0: $P(X \leq 0) = (.88)^4 = 0.5997 \geq 0.5, P(X < 0) = 0 \leq 0.5$
Q1 is 0: $P(X \leq 0) = (.88)^4 = 0.5997 \geq 0.25, P(X < 0) = 0 \leq 0.25$
Q3 is 1: $P(X \leq 1) = (.88)^4 + 4(.12)(.88)^3 = 0.5997 + 0.327 = 0.927 \geq 0.75, P(X < 1) = 0.5887 \leq 0.75$
IQR = Q3 - Q1 = 1

(c) $\overline{X} = 4/6 = 0.67, s = \sqrt{\frac{3(0-4/6)^2+2(1-4/6)^2+(2-4/6)^2}{5}} = \sqrt{0.67} = 0.816$

(d) Median is [0,1]: 50% of the data are less than or equal to zero, 0% are less than zero; 83.3% of the data are less than or equal to 1, 50% are less than one.
Q1 is 0
Q3 is 1
IQR = Q3 - Q1 = 1

(e) There are actually 24 cars in this study. 4 of them didn't pass inspection. $4/24 = 0.167$ is our best estimate for the probability that a randomly selected car doesn't pass inspection.