Recitation assignment due 3-2-10.
Partially covered for the review in lecture 3-1-10.
There will be a short bonus quiz in lecture 3-1-10.
Reminder: Exam 2 is 3-3-10.

z-table.

1. Use the z-table to verify or determine

a. $P(Z < -1.96) = 0.0250$ d. $P(Z > 2.65) =$

b. $P(Z < 1.96) = 0.9750$ e. $P(Z < - 2.65) =$

c. $P(| Z | < 1.96) = 0.9500$ f. $P(| Z | < 2.65) =$

g. Find z with $P(Z > z) = 0.021$ use closest table entry

h. Find z with $P(Z < z) = 0.021$ use closest table entry

i. Find z with $P(| Z | > z) = 0.042$ use closest table entry

z-approximation of Binomial X.

2. $n = 40$, $p = 0.4$

a. $E X = n p =$

b. $\text{Var } X = n p q =$

c. $\text{sd } X = \sqrt{n p q} =$

d. $P(X < 25) \sim P(Z < \frac{25 - n p}{\sqrt{n p q}}) =$

e. using continuity correction $P(X < 25) \sim P(Z < \frac{24.5 - n p}{\sqrt{n p q}}) =$

f. $P(X \geqslant 25) \sim$ use complement of (e)

g. check the two "at least 10" conditions

z-approximation of $\hat{p}$.

3. n = 40, p = 0.4

a. E $\hat{p}$ = $\frac{n\,p}{n}$ = p = (compare with 2a)

b. Var $\hat{p}$ = $\frac{n\,p\,q}{n^2}$ = $\frac{p\,q}{n}$ = (compare with 2b)

c. sd $\hat{p}$ = $\sqrt{\frac{p\,q}{n}}$ = (compare with 2c)

d. P($\hat{p} < \frac{25}{40}$) ~ P(Z < $\frac{\frac{25}{40} - p}{\sqrt{\frac{p\,q}{n}}}$) = P(Z < $\frac{\frac{25}{40} - 0.4}{\sqrt{\frac{0.4\;0.6}{40}}}$) =

same as (2d) ~ P(Z < $\frac{24.5 - n\,p}{\sqrt{n\,p\,q}}$) =

z-based approximate confidence interval for p.

4. n = 40, p = 0.4, X = 20 (p = probability of success, X = # successes in 40)

a. Determine z with P(| Z | < z) = 0.99 use closest table entry

b. 99% approx z-based CI for p is $\hat{p} \pm z \sqrt{\frac{\hat{p}\,\hat{q}}{n}}$ for z of 4a

note use of $\hat{p}$ in the estimate of sd of $\hat{p}$

c. What is the approximate probability that a sample of 40 would produce a 99% CI covering the actual p of 0.4?

d. Has the interval 4b covered the actual value of p = 0.4?

e. Agresti-Coull method is claimed to have coverage probability more nearly approximating the intended 99% for "most" n , p. The interval proceeds as above but artificially increases X to 22 and n to 44 (same as increasing the count of successes and failures each by 2). Determine the A-C 99% CI.

f. If the samples are withOUT-replacement rather than with-replacement the CI 4d has to be modified by FPC = $\sqrt{(N-n)/(N-1)}$ where N denotes the population size. Do so assuming a population of size 500. Has it made much difference?

g. Regardless of contexts, if 100,000 researchers independently prepare 99% CI (for their respective p) around how many of them will (unwittingly) have their CI fail to cover their p?

z-based approximate tests for p.

5. n = 100, H_0: p = 0.3, H_A: p > 0.3, X = # of successes in sample of n = 100.

a. Is the test 1-sided or 2-sided?

b. Suppose the test is set up to reject H_0 if $X \geqslant c$ for c = 37. Approximate the probability of type 1 error P($X \geqslant 37$ if p = 0.3) ~ P($Z > \frac{37 - 100\ 0.3}{\sqrt{100\ 0.3\ 0.7}}$).

c. Type 1 error probability is which one of the following?
   probability of rejecting the null hypothesis when it is true
   probability of failing to reject the null hypothesis when it is not true

d. Instead of setting up the test as in 5b let's just suppose we observe 37 successes in 100 trials. For this data the P-value is

$$P(X \geqslant 37 \text{ if } p = 0.3) \sim P\left(Z > \frac{37 - 100\ 0.3}{\sqrt{100\ 0.3\ 0.7}}\right).$$

e. Unrelated to the above, a journal has adopted the policy of rejecting H_0 : "worthless med" in favor of H_A : "med has beneficial effect" whenever data gives a P-value < 0.00001. Out of 100,000 worthless articles submitted to the journal (i.e. articles in which the null hypothesis of "worthless med" is actually true) around how many of the 100,000 worthless submissions will pass the journal's criterion? Note: that alone does not guarantee publication.

6. It is desired to set up a test of the hypothesis $p = p_0 = 0.02$ versus the alternative hypothesis that $p = p_1 = 0.03$ where p is the probability that a package will fail a burst test. It is desired that

type 1 error probability = P(we decide for p_1 if $p = p_0$ is true) = 0.01

type 2 error probability = P(we decide for p_0 if $p = p_1$ is true) = 0.02

Denote by X the number failing the burst test in a sample of n (n to be determined).

a. Determine z_0 with $P(Z > z_0) = 0.01$ using the closest table entry.

b. Determine z_1 with $P(Z < z_1) = 0.02$ using the closest table entry.

c. Determine the needed sample size n using the appropriate formula.

d. Determine the value c for which the test "reject the hypothesis $p = p_0 = 0.02$ if $X \geq c$" meets the above criteria. Use the appropriate formula.

e. Check that n from 1c meets the "at least 10" criterion for each of p_0 and p_1. If so, the normal approximation conditions that underlay the formula are met.

f. This test is good at discriminating between p = 0.2 and 0.3 or values of p to either side of these (David vs Goliath). The test will have trouble dealing with p values between 0.02 and 0.03 since the error probabilities were not controlled for there (in fact we did not cleave the line into null and alternative hypotheses at all).

formula sheet on exam (in addition to z-tables).

$\sqrt{(N - n)/(N - 1)}$

$\sqrt{n\,p\,q}$

$\sqrt{p\,q\,/\,n}$

$\sqrt{\hat{p}\,\hat{q}\,/\,n}$

$$n = \left(\frac{\sqrt{p_0\,q_0}\;\left|z_0\right| + \sqrt{p_1\,q_1}\;\left|z_1\right|}{p_0 - p_1}\right)^2$$

(on exam 2 the "n =" and "c =" will not be shown)

$$c = z_0\,\sqrt{n\,p_0\,q_0} + 0.5 + n\,p_0$$