Statistics 641, Fall 2012
Take Home Final Exam
Due Noon, Dec 22, 2012

1. Suppose that we have a binomial response with expected failure probabilities $\pi_0 = 0.2$ and $\pi_1 = 0.1$ for treatments 0 (control) and 1 (experimental) respectively. Calculate the required sample sizes to achieve 90% power (type II error rate of 10%) at level $\alpha = .01$ assuming

   (a) Equal allocation to groups 0 and 1

   (b) 2:1 randomization to groups 0, 1 respectively

   (c) 2:1 randomization to groups 1, 0 respectively

   (d) Why are these different?

2. Suppose that we are using minimization as a covariate adaptive allocation scheme and we wish to balance with respect to smoking status and sex. Using the notation from class let $G_t = |x_{11}^t - x_{12}^t| + |x_{21}^t - x_{22}^t|$.

   Suppose that the next subject is a non-smoking female and we have:

	Smoker		Sex		
Group	Y	N	M	F	Total
1	15	26	19	22	41
2	16	28	21	23	44

   To which treatment group should the next subject be allocated?

3. A randomized, two-arm trial is conducted comparing a control treatment (A) to a experimental treatment (B). The primary outcome is all-cause mortality.

   (a) After completion of the trial (between 1.5 and 3 years follow-up), we observe the following table:

	Subjects	Deaths	Person-years follow-up
A	500	241	835
B	500	219	876

   Assume exponential survival compute the hazard ratio for the difference between treatment groups.

   (b) It is noted that many subjects do not adhere to their assigned treatment and when subjects are classified by their level of adherence, we observe the following table:

	Adherence	Subjects	Deaths	Person-years follow-up
A	$> 80\%$	110	52	180
	50%–80%	240	153	352
	$\leq 50\%$	150	36	303
B	$> 80\%$	145	51	262
	50%–80%	280	148	463
	$\leq 50\%$	75	20	151

   Compute hazard ratios for subjects within each stratum based on adherence ($> 80\%$, 50%–80%, $\leq 50\%$), and note the differences with the overall comparison in part (a). Which result is more credible as an assessment of the effect of treatment and why?

4. Suppose we conduct a single arm trial with 30 subjects and a binomial outcome. Letting π be the true success probability, we wish to test the (one-sided) null hypothesis $H_0\colon \pi = 0.4$ versus $H_1\colon \pi > 0.4$ at a one-sided $\alpha = .05$.

   (a) In a fixed sample trial, we reject H_0 if, y, the number of successes out of 30 subjects, is larger than some integer b. Find the smallest b such that $\Pr\{y \geq b\} \leq 0.05$.

   (b) Suppose further that we perform 2 interim analyses when 10 and 20 subjects have completed, and a final analysis when all 30 subjects have completed. If y_k is the total number of successes observed at stage $k, k = 1, 2, 3$, we reject H_0 if $y_k \geq b_k$ for critical values $b_k, k = 1, 2, 3$. Using the α-spending function $g(t) = \alpha t$, we want to compute the critical values b_1, b_2, b_3 so that the cumulative rejection probability through stage k is at most $g(t_k)$, for $t_k = k/3$. (Note that because the data are discrete, we can't achieve exactly $g(t_k)$, so b_k will be the *smallest* critical value so that the rejection probabilities are at most $g(t_k)$)

   Find b_1 (at 10 subjects) using the exact binomial distribution of y_1 so that $\Pr\{y_1 \geq b_1\} \leq g(t_1)$

   (c) Because we may stop at stage 1, the sub-distribution, $f_2()$, of y_2 will not be binomial(20, .4).

      i. Compute $f_2(a) = \Pr\{y_2 = a\}$ for $a = 11, 12, 13, 14, 15$.

      ii. Compute $\Pr\{y_2 \geq a\}$ for $a = 11, 12, 13, 14, 15$.

      iii. Choose b_2 to be the smallest value of a for which $\Pr\{y_1 \geq b_1\} + \Pr\{y_2 \geq a\} \leq g(t_2)$.

   (d) Because we may stop at stages 1 and 2, the sub-distribution of y_3 will not be binomial(30, .4).

      i. Compute the $\Pr\{y_3 \geq a\}$ for $a = 16, 17, 18, 19, 20$ (note that for this part you will need to calculate $f_2(y)$ for all $y = 0, 1, \ldots, b_k - 1$).

      ii. Choose b_3 to be the smallest value of a for which $\Pr\{y_1 \geq b_1\} + \Pr\{y_2 \geq b_2\} + \Pr\{y_3 \geq a\} \leq g(1)$

   Compare b_3 to the critical value b from (a).

Notes: By using the `R` *functions* `sum`, `dbinom` *and* `pbinom` *and using vectors as in the homework 4 solutions, the computations are fairly straightforward. For (c), if you store the values of f_2 in a vector, it will be easier to do the calculations for part (d).*

These calculations are, in principle, identical to those for α-spending in the normal case. The difference is that the normal case requires that we perform complicated numerical integration. Because the data in this problem are discrete, instead of integrals are replaced by sums, which can be performed easily without specialized software. The purpose of this problem is to illustrate the mechanics of α-spending in a computationally simple way.

5. Suppose we conduct a trial with 3 dose groups (low, medium, high) and placebo (none), and a binary outcome (dead/alive). At the conclusion of the trial we observed the following table:

Treatment:	placebo ($i = 0$)	low ($i = 1$)	medium ($i = 2$)	high ($i = 3$)
Alive	10	18	22	25
Dead	90	82	78	75
	100	100	100	100

(a) Letting $\pi_i, i = 0, 1, 2, 3$ be the probability of death in group i (where i is defined in the table), the global null hypothesis is

$$H_{0123}\colon \pi_0 = \pi_1 = \pi_2 = \pi_3.$$

By reversing the roles of outcome and treatment, we may test this null hypothesis using the Wilcoxon rank-sum test (uncorrected). (This test is sensitive to the alternative hypothesis in which subjects who die are, for example, more likely to have received a lower dose than subjects who survive.)

Conduct this test of H_{0123} at level $\alpha = 0.05$.

(For convenience, the dataset `dataF.csv` has one record per subject with a treatment variable, `z`, and death variable `dead`. You can perform the Wilcoxon test directly using this dataset)

(b) We can implement a closed testing procedure as follows. First test the global null hypothesis H_{0123} at level $\alpha = .05$. If we reject H_{0123}, then test each of $H_{012}, H_{013}, H_{023}$, and H_{123} where $H_{012}\colon \pi_0 = \pi_1 = \pi_2$, $H_{013}\colon \pi_0 = \pi_1 = \pi_3$, etc, at level $\alpha = .05$ using the same test restricted to the treatment groups $\{0, 1, 2\}$, $\{0, 1, 3\}$, etc. (Note that the function `Wilcox.test` takes a `subset` argument, e.g., `subset=z!=3` to use only groups 0,1,2.)

Depending upon which hypotheses H_{jkl} are rejected, we can proceed to test pairwise hypotheses $H_{jk}\colon \pi_j = \pi_k$ at level $\alpha = 0.05$. (Note that for 2 groups and 2 outcomes, the Wilcoxon test (uncorrected) and the Pearson chi-square test (uncorrected) give identical results, so either can be used.)

Draw the graph (similar to the graphs in the Dec 11 lecture) showing the relationships between all 11 hypotheses and using this procedure, determine which of the pairwise hypotheses, H_{jk}, can be rejected at level $\alpha = 0.05$.

(c) Using the Bonferroni procedure for all 6 pairwise hypotheses, which can be rejected at overall level $\alpha = 0.05$?

(d) Using the Bonferroni procedure for the three pairwise hypotheses that compare an active dose to placebo, which can be rejected at overall level $\alpha = 0.05$?

(e) Using the Holm procedure for all 6 pairwise hypotheses, which can be rejected at overall level $\alpha = 0.05$?

(f) Using the Holm procedure for the three pairwise hypotheses that compare an active dose to placebo, which can be rejected at overall level $\alpha = 0.05$?