c681: Sample Questions / guidance on types of questions to be asked
{Wednesday, feb 24}

1. What distinguishes the Wilson CI for a proportion from the 'usual' large-sample CI taught in most elementary textbooks, and from the Klopper-Pearson CI, is the fact that it..

Circle the correct answer

(a) uses a continuity correction
(b) uses the Gaussian distribution, along with two binomial-based SD's, one calculated at the lower limit of π, one at the upper limit of π.
(c) uses the binomial probabilities to calculate two tail areas-- one based on the upper and one based on the lower limit
(d) uses a SE calculated at the point estimate
(e) uses a SD calculated at the null

2 The sampling variance of the log of a crude odds ratio calculated from a prevalence or 'risk' study is

$$1/c_1 + 1/c_0 + 1/nc_1 + 1/nc_0$$

where the subscripts 0 and 1 refer to exposed and unexposed, 'c' is the number of cases ("positives", individuals with the response/state/trait of interest), and 'nc' is the number of 'non-cases' ("negatives").

Regroup the terms in this variance formula into 2 parts so it is clear that it reflects the fact that (i) the variance of a difference of two independent statistics is the sum of their variances and (ii) the log of a ratio is the difference of the logs of the components

3 Questions 2, 3 and 4 of Assignment 2

A number of students have told me they are 'lost at sea' in relation to the Poisson distribution. The question on the exam will be similar to parts of Questions 2, 3 and 4 of Assignment 2

4 Many people with sleep apnea unintentionally fall asleep during the day. In a study of how long such people stay asleep, subjects were brought into an observation room in the afternoon. A research assistant ("RA-1") noted the time at which the subjects began to sleep and measured the duration that they slept (both easier said that done, but put these technical matters aside). At 5 pm, the research assistant went home. If a subject was sleeping at that time, the RA recorded that fact in his log, but left the subject sleeping, and told the security person "when the person wakes up, just tell her/him (s)he can leave now."

In the same sleep study, a different RA ("RA-2") watched until the subject fell asleep, then checked back on the subject every 15 minutes, and if a subject was sleeping when the RA left for the day, she told the security person to note the time and the subject 's status (still sleeping, awake) when the security person did his rounds of the building every 30 minutes.

Yet another RA ("RA-3") observed all the subjects from the beginning, but sometimes fell asleep himself. When the RA woke up, he noted that some patients, who had not yet fallen asleep before the RA himself dozed off, were now sleeping. He noted the time that he (the RA) woke up, and the fact that the subject was already sleeping, and then continued to observe until the subject woke up.

In a study of students and faculty, another RA ("RA-4") quietly entered the back of room 25 of Purvis Hall at 4.30 each Monday (in the middle of the epidemiology seminar) and noted who in the audience was sleeping at that time (not easy to do from the back of the room, but again, leave this technical aspect aside). The RA observed how long until these people woke up.

(a) *Describe the kinds of observations produced by the 4 different RA's (distinguish the various types of 'censoring'; identify cases where there might be 'truncation')*

(b) *From the data collected by RA-1, is it possible to estimate the distribution (or part of the distribution) of the duration spent sleeping? If so, indicate how; if not, say why not.*

(c) *Even if RA-4 were able to find out when his sample of sleepers began sleeping, what biases would he create if he analyzed these observations as if they came from a simple random sample of 'seminar sleepers'?*

(d) Are the data collected by RA-3 subject to the same problem as those of RA-4? .Justify your answer.

There will be a (less complex) example of censored sleep-duration data on the exam.

5 You are given excerpts from a current lifetable [just *l, d and e*) and asked questions about it.

6 In connection with the trend analysis of injury rates in assignment #7, a student e-mailed JH to say

"My fitted equation was

```
logit[ proportion injured | year ] = -4.8  -0.0032 × year "
```

"When I convert the logit to a proportion for year = 0 (beta = -0.0032), I get pi=0.5. How can the probability of injury increase by 0.5 for every unit of time (1-year) when pi has to be between 0 and 1? "

Answer this student. Hint: what is the biggest difference between beta_0 and beta_1? the word difference (and its transformed counterpart) is part of the answer.

7 I will ask questions based on the following excerpts from the abstract of the report "Randomized controlled trial of labouring in water compared with standard of augmentation for management of dystocia [slow or difficult labor] in first stage of labour" and on a set of regression analyses I will supply in the exam.

> To evaluate the impact of labouring in water during first stage of labour on rates of epidural analgesia and operative delivery in nulliparous women with dystocia, 99 nulliparous women with dystocia (cervical dilation rate < 1 cm/hour in active labour) at low risk of complications were randomized to immersion in water in birth pool or standard augmentation for dystocia (amniotomy and intravenous oxytocin).
>
> Women randomised to immersion in water had a lower rate of epidural analgesia than women allocated to augmentation (47% v 66%, relative risk 0.71 (95% confidence interval 0.49 to 1.01), number needed to treat for benefit (NNT) 5).
>
> Six neonates born to women in the water labour group were admitted to the neonatal unit compared with none in the augmentation group (P =0.013).

Some of the questions on 7. will be similar to some of those in Q1 and Q3 of assignment #5

8 I will give a computer printout and ask questions about 1 study similar to

- Leprosy example (on exam web page). .
- Down's syndrome example (assignment #7).
- Autism example (assignment #3 and #7). I will give a printout and ask questions about it.

Based on the study, I may also pose some questions similar to those on assignment #3 (but not the Robins-Greenland-Breslow CI)

9 In assignment 5, you were asked to calculate the probability of capsule invasion for a given PSA; but what if the question had been: at what PSA level is the fitted probability = 25%? = 75%

10 You were given the sex-specific equations for ideal weight as a function of height. Be able to write these as a single master equation, and able to go backwards and forwards between the equation and a graph.