Pharmacy 2012 - Winter 2020 - Biostatistics - Assignment 3 solutions

(out of 40 points)

1. (10 points) Prostate-specific antigen (PSA) is a tissue-specific tumor marker used by urologists and oncologists to monitor treatment responses, prognosis, and progression in patients with prostatic cancer. A high level of PSA in the blood is indicative of prostate cancer. Researchers have found that the sensitivity of the test is .80, and that there is a .25 probability of testing positive if you don't have prostate cancer. The prevalence of prostate cancer in a particular population of males is 15%.

   We are given $P(T^+|D^+) = .80$, $P(T^+|D^-) = .25$ and $P(D+) = .15$.

   The most convenient way to carry out the calculations is to first calculate the joint probabilities $P(T^+ \cap D^+)$, $P(T^+ \cap D^-)$, $P(T^- \cap D^+)$, and $P(T^- \cap D^-)$, combine them to get any required marginal probabilitites, and then calculate conditional proabilities, as needed.

$$P(T^+ \cap D^+) = P(T^+|D^+)P(D^+) = .8(.15) = .12$$

$$P(D^-) = 1 - P(D^+) = 1 - .15 = .85$$

$$P(T^+ \cap D^-) = P(T^+|D^-)P(D^-) = .25(.85) = .2125$$

$$P(T^-|D^-) = 1 - P(T^+|D^-) = 1 - .25 = .75$$

$$P(T^- \cap D^-) = P(T^-|D^-)P(D^-) = .75(.85) = .6375$$

$$P(T^-|D^+) = 1 - P(T^+|D^+) = 1 - .80 = .20$$

$$P(T^- \cap D^+) = P(T^-|D^+)P(D^+) = .20(.15) = .03$$

   From these, we find:

(1) (a) What is the probability of False Negative of the test?

$$P(FN) = 1 - P(T^+|D^+) = 1 - .8 = .2$$

(1) (b) What is the probability of False Positive of the test?

$$P(FP) = 1 - P(T^+|D^-) = .25$$

(2) (c) What is the specificity of the test? $P(T^-|D^-) = .75$

(2) (d) The probability of testing positive is $P(T^+) = P(T^+ \cap D^+) + P(T^+ \cap D^-) = .12 + .2125 = .3325$.

(2) (e) What is the probability of a subject actually having prostate cancer if they have a positive result on the PSA test?
$P(D^+|T^+) = P(D^+ \cap T^+)/P(T^+) = .12/.3325 \approx .3609$.

(2) (f) What is the probability that a subject does not have the disease if they have a negative PSA test?
$P(D^-|T^-) = \frac{P(D^- \cap T^-)}{P(T^-)} = \frac{.6375}{1-.3325} \approx .955$

2. (7 points) The Physicians' Health Study was a randomized double-blind placebo- controlled trial of beta-carotene (50mg every other day). 22,071 male physicians age 40-84 were enrolled in 1982. The subjects were followed until December 1995, for the development of new cancers (malignant neoplasms). The group of size 11,036 receiving beta-carotene had 1273 new cancers, while the placebo group of size 11,035 had 1293 new cancers.

(2) (a) What is the estimated relative risk for malignant neoplasms?

$\frac{1273/11036}{1293/11035} \approx .984$

(5) (b) Calculate the 95% confidence interval for the relative risk

The estimated log relative risk is $\approx log(.984) \approx -.016$,

with estimated standard error $\sqrt{1/1273 - 1/11036 + 1/1293 - 1/11035} \approx .037$.

The 95% confidence interval for the log(RR) is $-.016 \pm 1.96(.037)$, or $-.016 \pm .073$, or (-.089, .057).

The 95% confidence interval for the RR is $(e^{-.089}, e^{.057})$ or approximately (.92,1.06).

(Subtract only 1 point if there was an error in the standard error, but the calculations were otherwise correct. If the relative risk in part (a) was wrong, don't subtract extra points for an incorrect log(RR) in part (b). Otherwise, subtract 1 point for each error.)

3. (8 points) A retrospective study obtained the following results in 168 patients admitted to hospital with perforated peptic ulcer and a set of matched controls.

	NSAID Use		
	Yes	No	Total
Cases	79	89	168
Controls	12	156	168

(2) (a) Calculate the odds ratio for the association between perforated peptic ulcer and nonsteroidal anti-inflammatory drugs (NSAID) use, in the Yes relative to the No group. $(79/89)/(12/156) \approx 11.54$.

(4) (b) Calculate the 95% confidence interval for the odds ratio.

The estimated log odds ratio is $\approx log(11.54) \approx 2.446$,

with estimated standard error $\sqrt{1/79 + 1/89 + 1/12 + 1/156} \approx .337$

The 95% confidence interval for the log odds ratio is $2.446 \pm 1.96(.337)$, or $2.446 \pm .661$, or (1.786, 3.107).

The confidence interval for the odds ratio is, approximately $(e^{1.786}, e^{3.107})$, or (5.97,22.35).

(subtract 1 point for each error).

(2) (c) Based on this confidence interval, is there significant evidence against the null hypothesis of no association between NSAID use and peptic ulcers at the $\alpha = .05$ level of significance? Explain.

Yes, because the 95% CI for the odds ratio does NOT contain 1.

	Therapy			
Healed	R	L15	L30	Total
Yes	61	81	85	227
No	54	37	32	123
Total	115	118	117	350

4. (15 points) In a study on the effect of lansoprazole (15mg and 30 mg) and ranitidine on ulcer healing, Agrawal et al obtained the numbers in the table above.

   Do an overall test to determine whether there are any differences among the therapies, in the following steps.

(2) (a) State the hypotheses.

one of

- H0: no association between therapy and healing, HA: there is an association
  OR
- H0: therapy and healing are independent, HA: therapy and healing not independent
  OR
- H0: row distributions are the same, HA: row distributions not the same
  OR
- H0: column distributions are the same, HA: column distributions are different

(4) (b) Calculate the expected counts, and present the results in a table like the following.

Expected counts

	Therapy		
Healed	R	L15	L30
Yes	74.59	76.53	75.88
No	40.41	41.47	41.12

For example, the (1,1) entry is, 227(115)/350 ≈ 74.59.

(subtract 1 point for each error)

(4) (c) Calculate the contributions to the goodness of fit test statistic, and present the results in a table like the following.

Contributions to the goodness of fit statistic

	Therapy		
Healed	R	L15	L30
Yes	2.47	0.26	1.10
No	4.57	0.48	2.02

The (1,1) entry is, for example, $(61 - 74.59)^2/74.59 \approx 2.47$.

(subtract 1 point for each error)

(2) (d) Calculate the observed value of the test statistic. The sum of the entries in the previous table, or ≈ 10.90

(1) (e) What is the degrees of freedom? $(2-1)(3-1) = 2$

(2) (f) Determine the P value as accurately as possible using the χ^2 table on the class web site.

$P(\chi^2_2 > 10.90) < .005$