Recitation 4 - P8130 Fall 2017

October 6, 2017

Problem 1

Rosner textbook 8^{th} edition, 7.72-7.74.
Breast cancer is strongly influenced by a woman's reproductive history. In particular, the longer the length of time from the age at menarche (the age when menstruation begins) to the age at first childbirth, the greater the risk is for breast cancer. A projection was made based on a mathematical model that the 30-year risk of a woman in the general U.S. population developing breast cancer from age 40 to age 70 is 7%. Suppose a special subgroup of five hundred 40-year-old women without breast cancer was studied whose age at menarche was 17 (compared with an average age at menarche of 13 in the general population) and age at first birth was 20 (compared with an average age at first birth of 25 in the general population). These women were followed for development of breast cancer between ages 40 and 70. The study found that 18 of the women developed breast cancer between age 40 and age 70.

7.72 Test the hypothesis that the underlying rate of breast cancer is the same or different in this group as in the general population.

7.73 Provide a 95% CI for the true incidence rate of breast cancer over the period from age 40 to 70 in this special subgroup.

7.74 Suppose 100 million women in the U.S. population have not developed breast cancer by the age of 40. What is your best estimate of the number of breast-cancer cases that would be prevented from age 40 to 70 if all women in the U.S. population reached menarche at age 17 and gave birth to their first child at age 20? Provide a 95% CI for the number of breast-cancer cases prevented.

Problem 2

Many children have tympanostomy tubes surgically inserted in their ears to reduce hearing loss associated with persistent otitis media and prevent recurrences of episodes of otitis media after tubes are inserted. However, acute otorrhea (a discharge from the external ear indicating inflammation of the external or middle ear), where middle ear fluid drains through the tube, is a common side effect with tympanostomy tubes.
A clinical trial was conducted (Van Dongen et al.) among children 1–10 years of age with prior symptoms of otorrhea comparing efficacy of (i) antibiotic eardrops, (ii) oral antibiotics, and (iii) observation without treatment, referred to below as observation. Children were seen at home by study physicians at 2 weeks and 6 months after randomization. The primary outcome was the presence of otorrhea at 2 weeks observed by study physicians. The results are given in Table 1.

Table 1: Number of children with otorrhea at 2 weeks of follow-up

Group	Number of children	Number of children with otorrhea at 2 weeks
Antibiotic ear drops	76	12
Oral antibiotics	77	34
Observation	75	41

1. Provide a point estimate and a 95% confidence interval for the prevalence of otorrhea at 2 weeks in the observation group (i.e. the proportion of children with otorrhea at 2 weeks among all children in the observation group).
2. Provide a point estimate and a 95% CI for the prevalence of otorrhea in the ear drop group.
3. What test can be used to compare the prevalence of otorrhea for the ear drop group vs. the observation group? State the hypotheses to be tested.
4. Perform the test in question 3 and report a p-value (two-tailed). Interpret your results in words.

Problem 3

A study was performed comparing the mean ERG amplitude of patients with different genetic types of retinitis pigmentosa (RP), a genetic eye disease that often results in blindness. The results obtained for ln(ERG amplitude) among patients 18-29 years of age, are shown below.

	Genetic Type	Mean ± SD	n
1	Dominant	0.91 ± 0.15	66
2	Recessive	0.34 ± 0.23	34
3	Sex-Linked	-0.07 ± 0.23	25

a. Use an appropriate hypothesis test to assess if there are overall group differences in mean ln(ERG amplitude) by genetic type. Construct an ANOVA table.

b. Assess if there are differences between each pair of genetic types using appropriate testing methods.

Problem 4

Using the SAS output below to answer the following questions:

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	22.671931	7.557310	2.64	0.0492
Error	397	1136.797016	2.863469		
Corrected Total	400	1159.468947			

R-Square	Coeff Var	Root MSE	viral Mean
0.019554	50.61496	1.692179	3.343238

Source	DF	Anova SS	Mean Square	F Value	Pr > F
trtarm	3	22.67193098	7.55731033	2.64	0.0492

Comparisons significant at the 0.05 level are indicated by ***.				
trtarm Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
D - C	0.3444	-0.2181	0.9069	
D - B	0.4864	-0.2113	1.1842	
D - A	0.6127	-0.0082	1.2335	
C - D	-0.3444	-0.9069	0.2181	
C - B	0.1420	-0.5655	0.8495	
C - A	0.2683	-0.3635	0.9000	
B - D	-0.4864	-1.1842	0.2113	
B - C	-0.1420	-0.8495	0.5655	
B - A	0.1262	-0.6285	0.8810	
A - D	-0.6127	-1.2335	0.0082	
A - C	-0.2683	-0.9000	0.3635	
A - B	-0.1262	-0.8810	0.6285	

a. Suppose researchers wish to test if viral load at 12 months differs by treatment arm. Perform an appropriate test at the 5% level of significance.

b. Are pairwise hypothesis tests appropriate here? If so, which groups are significantly different from one another? What alpha level should be used for each pair-wise test?