

Homework 4 (Due on Feb 13)

February 6, 2014

STA 5172

Problem I

Studies have been undertaken to assess the relationship between abortion and the development of breast cancer. In one study among nurses (the Nurses Health Study II), there were 16,359 abortions among 2,169,321 person-years of follow-up for women of reproductive age. (Note: 1 person-year = 1 woman followed for 1 year.)

1. What is the expected number of abortions among nurses over this time period if the national incidence of abortion is 25 per 1000 women per year and no woman has more than 1 abortion?
2. Does the abortion rate among nurses differ significantly from the national experience? Why or why not? (Test based on 5% level of significance)

Problem II

The data in Table 6.8 concern the mean triceps skin-fold thickness in a group of normal men and a group of men with chronic airflow limitation.

Table 6.8 Triceps skin-fold thickness in normal men and men with chronic airflow limitation

Group	Mean	<i>sd</i>	<i>n</i>
Normal	1.35	0.5	40
Chronic airflow limitation	0.92	0.4	32

Source: Reprinted with permission of *Chest*, 85(6), 58S–59S, 1984.

1. What is the standard error of the mean for each group?
2. Find the upper 1st percentile of a t distribution with 16 df.
3. Find the lower 10th percentile of a t distribution with 28 df.
4. Find the upper 2.5th percentile of a t distribution with 7 df.

5. What are the upper and lower 2.5th percentiles for a chi-square distribution with 2 df? What notation is used to denote these percentiles?

Problem III

Consider the hospital stay data below. Also find the link in the webpage

<http://www.stat.fsu.edu/~debdeep/hospital.txt>

1. Compute a 95% CI for the mean duration of hospitalization.
2. Compute a 95% CI for the mean white blood count following admission.
3. Answer Problem 2 for a 90% CI.
4. What is the relationship between your answers to Problems 2 and 3?

Problem IV

Suppose 100 hypertensive people are given an antihypertensive drug and the drug is effective in 20 of them. By effective, we mean their DBP is lowered by at least 10 mm Hg as judged from a repeat blood-pressure measurement 1 month after taking the drug.

1. What is the best point estimate of the probability p of the drug being effective?
2. *A placebo is a simulated or otherwise medically ineffectual treatment for a disease or other medical condition intended to deceive the recipient. Sometimes patients given a placebo treatment will have a perceived or actual improvement in a medical condition, a phenomenon commonly called the placebo effect.* Suppose we know that 10% of all hypertensive patients who are given a placebo will have their DBP lowered by 10 mm Hg after 1 month. Can we carry out some procedure to be sure we are not simply observing the placebo effect?
3. What assumptions have you made to carry out the procedure in the previous question?
4. Suppose we decide a better measure of the effectiveness of the drug is the mean decrease in blood pressure rather than the measure of effectiveness used previously. Let $d_i = x_i - y_i, i = 1, \dots, 100$ where x_i = DBP for the i th person before taking the drug and y_i = DBP for the i th person 1 month after taking the drug. Suppose the sample mean of the d_i is +5.3 and the sample variance is 144.0.

Table 2.11 Hospital-stay data

ID no.	Duration of hospital stay	Age	Sex 1 = M 2 = F	First temp. following admission	First WBC ($\times 10^3$) following admission	Received antibiotic 1 = yes 2 = no	Received bacterial culture 1 = yes 2 = no	Service 1 = med. 2 = surg.
1	5	30	2	99.0	8	2	2	1
2	10	73	2	98.0	5	2	1	1
3	6	40	2	99.0	12	2	2	2
4	11	47	2	98.2	4	2	2	2
5	5	25	2	98.5	11	2	2	2
6	14	82	1	96.8	6	1	2	2
7	30	60	1	99.5	8	1	1	1
8	11	56	2	98.6	7	2	2	1
9	17	43	2	98.0	7	2	2	1
10	3	50	1	98.0	12	2	1	2
11	9	59	2	97.6	7	2	1	1
12	3	4	1	97.8	3	2	2	2
13	8	22	2	99.5	11	1	2	2
14	8	33	2	98.4	14	1	1	2
15	5	20	2	98.4	11	2	1	2
16	5	32	1	99.0	9	2	2	2
17	7	36	1	99.2	6	1	2	2
18	4	69	1	98.0	6	2	2	2
19	3	47	1	97.0	5	1	2	1
20	7	22	1	98.2	6	2	2	2
21	9	11	1	98.2	10	2	2	2
22	11	19	1	98.6	14	1	2	2
23	11	67	2	97.6	4	2	2	1
24	9	43	2	98.6	5	2	2	2
25	4	41	2	98.0	5	2	2	1

5. What is the standard error of d ?
6. What is a 95% CI for the population mean of d ?
7. Can we make a statement about the effectiveness of the drug?
8. What does a 95% CI mean, in words, in this case?