Extra Exercises: Chapters 1–3
Statistics 301
Professor Wardrop
Spring 2008

Section 2.3

1. Consider a balanced study with eight subjects, identified as A, B, C, D, E, G, H, and J. In the actual study,

   - A, B, C and D are assigned to the first treatment, and
   - There are exactly four successes, and they are obtained by A, B, C, and H.

   This information is needed for parts (a)–(c) below.

   (a) Compute the observed value of the test statistic.

   (b) Assume that the Skeptic is correct. Determine the observed value of the test statistic for the assignment that places A, D, E, and G on the first treatment, and the remaining subjects on the second treatment.

   (c) We have obtained the sampling distribution of the test statistic on the assumption that the Skeptic is correct. It also is possible to obtain a sampling distribution of the test statistic if the Skeptic is wrong *provided* we specify *exactly* how the Skeptic is in error. Assume that the Skeptic is incorrect about subjects C, D, H, and J, but correct about subjects A, B, E, and G. This means that for subjects C, D, H, and J, his/her/its response will change if the treatment changes.

   For the assignment that puts A, D, E, and H on the first treatment, and the other subjects on the second treatment, determine the response for each of the eight subjects.

2. Consider a unbalanced study with nine subjects, identified as A, B, C, D, E, G, H, J, and K. In the actual study,

   - A, B, C, D, and E are assigned to the first treatment, and
   - There are exactly five successes, and they are obtained by B, C, E, H, and J.

   This information is needed for parts (a)–(c) below.

   (a) Compute the observed value of the test statistic.

   (b) Assume that the Skeptic is correct. Determine the observed value of the test statistic for the assignment that places A, C, D, G, and K on the first treatment, and the remaining subjects on the second treatment.

   (c) Assume that the Skeptic is incorrect about subjects C, D, E, G, H, J, and K, but correct about subjects A, and B.

   For the assignment that puts A, E, G, H, and J on the first treatment, and the other subjects on the second treatment, determine the response for each of the nine subjects.

3. An unbalanced yields the data below.

Treatment	S	F	Total
1	a	b	10
2	c	d	5
Total	10	5	15

   On the assumption the Skeptic is correct, list all possible values of the test statistic.

4. A comparative study yields the following data.

Treatment	S	F	Total
1	a	b	5
2	c	d	4
Total	6	3	9

   On the assumption the Skeptic is correct, determine all possible values of the test statistic.

Section 2.5

5. Below is the sampling distribution of the test statistic for Fisher's test for a comparative study.

x	$P(X = x)$	$P(X \leq x)$	$P(X \geq x)$
-0.405	0.0003	0.0003	1.0000
-0.340	0.0019	0.0022	0.9997
-0.275	0.0104	0.0126	0.9978
-0.210	0.0378	0.0504	0.9874
-0.145	0.0973	0.1477	0.9496
-0.080	0.1782	0.3259	0.8523
-0.015	0.2335	0.5594	0.6741
0.050	0.2172	0.7766	0.4406
0.115	0.1410	0.9176	0.2234
0.180	0.0620	0.9796	0.0824
0.245	0.0174	0.9970	0.0204
0.310	0.0028	0.9998	0.0030
0.375	0.0002	1.0000	0.0002

   (a) Find the P-value for the second alternative ($p_1 < p_2$) if:
      i. $x = -0.275$.
      ii. $x = -0.145$.
      iii. $x = -0.080$.

   (b) Find the P-value for the first alternative ($p_1 > p_2$) if:
      i. $x = -0.015$.
      ii. $x = 0.115$.
      iii. $x = 0.180$.

   (c) Determine the P-value for the third alternative ($p_1 \neq p_2$) for each of the x's given in (a).

   (d) Determine the P-value for the third alternative ($p_1 \neq p_2$) for each of the x's given in (b).

   (e) Determine the value of x and the P-value that satisfy the following condition: The data are statistically significant but not highly statistically significant for the first alternative ($p_1 > p_2$).

   (f) Determine the values (there are two) of x and the P-values (a different one for each x) that satisfy the following condition: The data are statistically significant but not highly statistically significant for the third alternative ($p_1 \neq p_2$).

6. Sally performs a comparative study with a dichotomous response. She obtains the sampling distribution of the test statistic for Fisher's test for her data; it is given below.

x	$P(X = x)$	$P(X \leq x)$	$P(X \geq x)$
-0.30	0.0003	0.0003	1.0000
-0.25	0.0029	0.0032	0.9997
-0.20	0.0151	0.0184	0.9968
-0.15	0.0514	0.0697	0.9816
-0.10	0.1193	0.1890	0.9303
-0.05	0.1957	0.3848	0.8110
0.00	0.2305	0.6152	0.6152
0.05	0.1957	0.8110	0.3848
0.10	0.1193	0.9303	0.1890
0.15	0.0514	0.9816	0.0697
0.20	0.0151	0.9968	0.0184
0.25	0.0029	0.9997	0.0032
0.30	0.0003	1.0000	0.0003

   (a) Find the P-value for the second alternative ($p_1 < p_2$) if:
      i. $x = -0.25$.
      ii. $x = -0.15$.
      iii. $x = -0.05$.

   (b) Find the P-value for the first alternative ($p_1 > p_2$) if:
      i. $x = 0.00$.
      ii. $x = 0.10$.
      iii. $x = 0.20$.

   (c) Determine the P-value for the third alternative ($p_1 \neq p_2$) for each of the x's given in (a).

   (d) Determine the P-value for the third alternative ($p_1 \neq p_2$) for each of the x's given in (b).

   (e) Determine the value of x and the P-value that satisfy the following condition: The data are statistically significant but not highly statistically significant for the first alternative ($p_1 > p_2$).

(f) Determine the values (there are two) of x and the P-value that satisfy the following condition: The data are statistically significant but not highly statistically significant for the third alternative ($p_1 \neq p_2$).

7. Two comparative studies with dichotomous responses, randomization, and two treatments are performed. The first study has $n = 35$ and the second study is balanced.

   The observed value of the test statistic, x, is a positive number for the first study and a negative number for the second study. The P-values for all three alternatives are obtained for both studies. The six (sorted) P-values are below.

   0.0433, 0.0866, 0.3297, 0.5060, 0.8700, and 0.9865.

   Match each P-value with its study and alternative.

Study	$>$	Alternative $<$	$\neq$
First			
Second			

8. Two comparative studies, with two treatments, dichotomous responses and randomization are performed. The first study is balanced.

   The observed value of the test statistic, x, is a positive number for the first study and a negative number for the second study. The P-values for all three alternatives are obtained for both studies. The six (sorted) P-values are below.

   0.1001, 0.1223, 0.2002, 0.2187, 0.9432, and 0.9565.

   Match each P-value with its study and alternative. **Hint: For the second study, for the actual x,**

$$P(X = x) = 0.0655.$$

Study	$>$	Alternative $<$	$\neq$
First			
Second			

9. A comparative study, with two treatments, dichotomous response and randomization is performed. The study is not balanced.

   The observed value of the test statistic, x, is a negative number. The exact P-values for all three alternatives are obtained and are below.

   0.3390, 0.5558 and 0.8466.

   Match each P-value with its alternative.

$>$	Alternative $<$	$\neq$

10. A comparative study, with two treatments, dichotomous response and randomization is performed. The study is not balanced.

The observed value of the test statistic, x, is a negative number. The exact P-values for all three alternatives are obtained and are below.

0.1026, 0.1902 and 0.9637.

Match each P-value with its alternative.

	Alternative	
$>$	$<$	$\neq$

11. A comparative study, with two treatments, dichotomous response and randomization is performed. The study is balanced.

The observed value of the test statistic, x, is a positive number. The exact P-values for all three alternatives are obtained. The three exact P-values are below along with three other numbers. Also note that for the actual value of x, $P(X = x) = 0.1337$.

0.0763, 0.1815, 0.4287, 0.7050, 0.8574 and 0.9307,

Identify the three P-value and match each with its alternative.

	Alternative	
$>$	$<$	$\neq$

12. A comparative study, with two treatments, dichotomous response and randomization is performed. The study is balanced.

The observed value of the test statistic, x, is a positive number. The exact P-values for all three alternatives are obtained. The three exact P-values are below along with three other numbers. Also note that for the actual value of x, $P(X = x) = 0.1084$.

0.0463, 0.2015, 0.4436, 0.6648, 0.8872 and 0.9207,

Identify the three P-value and match each with its alternative.

	Alternative	
$>$	$<$	$\neq$

Section 3.1

13. Below is the sampling distribution of the test statistic for Fisher's Test for an unbalanced comparative study.

x	$P(X = x)$	x	$P(X = x)$
-0.8	0.0070	0.1	0.3916
-0.5	0.0932	0.4	0.1632
-0.2	0.3263	0.7	0.0186

Draw the probability histogram for this distribution.

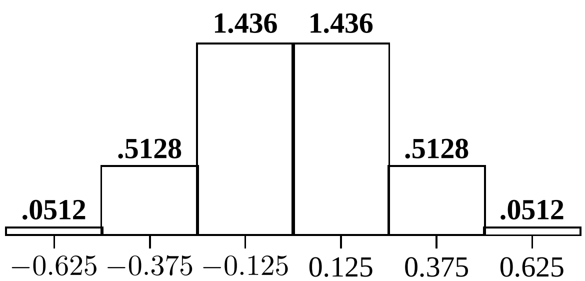
14. Below is the partial sampling distribution of the test statistic for Fisher's Test for a balanced comparative study.

x	$P(X = x)$	x	$P(X = x)$
-0.8	0.0004		
-0.6	0.0095		
	0.0750		
	0.2401		

   (a) Complete the table.

   (b) Draw the probability histogram for this distribution.

15. Below is the probability histogram for the test statistic for a Fisher's test. Obtain the sampling distribution and present it in a table (i.e. a column of x's and a column of probabilities).



   (a) Obtain the sampling distribution and present it in a table (i.e. a column of x's and a column of probabilities).

   (b) Given that the actual x equals 0.375, obtain the exact P-value for the alternative $p_1 > p_2$.

16. Maria draws a probability histogram for a Fisher's test. You are given the following facts:

   - The study is balanced.
   - There is a rectangle centered at 0.000 and a rectangle centered at 0.025. These rectangles "touch;" that is, there are no other rectangles between them.
   - The rectangle centered at 0.000 has a height of 5.784.

   (a) Calculate $P(X = 0.000)$.

   (b) Given that the actual value of the test statistic is $x = 0.025$, calculate the exact P-value for the alternative $p_1 > p_2$.

17. Nancy draws a probability histogram for a Fisher's test. You are given the following facts:

   - The study is balanced.
   - There is a rectangle centered at 0.00 and a rectangle centered at 0.08. These rectangles "touch;" that is, there are no other rectangles between them.
   - The rectangle centered at 0.00 has a height of 2.835.

   (a) Calculate $P(X = 0)$.

   (b) Given that the actual value of the test statistic is $x = 0.08$, calculate the P-value for the alternative $p_1 > p_2$.

Section 3.2

18. Below is the sampling distribution of the test statistic for Fisher's Test for an unbalanced comparative study.

x	$P(X = x)$	x	$P(X = x)$
-0.8	0.0070	0.1	0.3916
-0.5	0.0932	0.4	0.1632
-0.2	0.3263	0.7	0.0186

I performed a simulation experiment with 5,350 runs. The frequencies of occurrences of the six values of the test statistic were obtained and then sorted:

34, 107, 482, 868, 1769, and 2090.

Which of these five numbers is the observed frequency of $x = 0.4$? Briefly justify your answer.

19. Below is the sampling distribution of the test statistic for Fisher's Test for an unbalanced comparative study.

x	$P(X = x)$	x	$P(X = x)$
-0.8	0.0037	0.1	0.4396
-0.5	0.0732	0.4	0.1538
-0.2	0.3297		

I performed a simulation experiment with 6,250 runs. The frequencies of occurrence of the five values of the test statistic were obtained and then sorted:

17, 450, 975, 2073, and 2735.

Which of these five numbers is the observed frequency of $x = 0.4$? Briefly justify your answer.

20. A sample space has three possible outcomes, B, C, and D. It is known that $P(C) = P(D)$. The operation of the chance mechanism is simulated 10,000 times (runs). The sorted frequencies of the three outcomes (B, C, and D) are:

1973, 2021, and 6006.

(a) What is your approximation of $P(B)$?

(b) Present and justify three possible approximations of $P(C)$.

Section 3.4

21. An unbalanced comparative study has a total of 250 subjects, with 100 subjects on treatment 1. The total number of successes is 135, with 45 of the successes on the first treatment.

Use the standard normal curve to obtain an approximate P-value for Fisher's test with the third alternative ($\neq$).

22. An unbalanced comparative study has a total of 280 subjects, with 100 subjects on treatment 1. The total number of successes is 153, with 45 of the successes on the first treatment.

Use the standard normal curve to obtain an approximate P-value for Fisher's test with the third alternative ($\neq$).

23. An unbalanced comparative study has a total of 380 subjects, with 80 subjects on treatment 1. The total number of successes is 123, with 24 of the successes on the first treatment.

Use the standard normal curve to obtain an approximate P-value for Fisher's test with the second alternative ($<$).

24. A comparative study yields the following data.

Treatment	S	F	Total
1	200	200	400
2	330	270	600
Total	530	470	1000

(a) Use the standard normal curve to obtain an approximate P-value for Fisher's test with the first alternative ($>$).

(b) Use the standard normal curve to obtain an approximate P-value for Fisher's test with the second alternative ($<$).

(c) Use the standard normal curve to obtain an approximate P-value for Fisher's test with the third alternative ($\neq$).