MA 486/586-1C (Statistics), Dr Chernov | Midterm test #2

Each problem is 20 points (100 total). Show your work. | Fri, Mar 21, 2014

1. In a biology laboratory students test the Mendelian theory of inheritance using corn. The Mendelian theory claims that the frequencies of the four categories smooth and yellow, wrinkled and yellow, smooth and purple, and wrinkled and purple will occur in the ratio 9:3:3:1. If a student counted 341, 139, 112, and 48, respectively, for these four categories, would these data support the Mendelian theory? Let $\alpha = 0.10$.

Bonus: what can you say about the p-value of the test?

Solution:

$$Q = \frac{19^2}{360} + \frac{19^2}{120} + \frac{8^2}{120} + \frac{8^2}{40} = 6.144$$

The critical region is $Q > \chi^2_{0.1} = 6.251$. We accept H_0, supporting the Mendelian theory.

The p-value is 0.1048.

2. Suppose a sample $x_1, \dots, x_9$ from a normal distribution $N(\mu, \sigma^2)$ yielded $\bar{x} = -4.9$ and $s^2 = 36$. Assume $\alpha = 0.1$.

(a) Test the hypothesis $H_0 : \mu = -1$ against $H_1 : \mu \neq -1$.

(b) Test the hypothesis $H_0 : \sigma = 9$ against $H_1 : \sigma < 9$.

[Bonus] What can you say about the p-values in the above tests?

Solution:

(a) the test statistic is

$$T = \frac{-4.9 - (-1)}{6/\sqrt{9}} = -1.95$$

The critical region is $|T| > t_{0.05}(8) = 1.860$. We accept H_1.

The p-value is 0.087.

(b) the test statistic is

$$\chi^2 = \frac{8 \cdot 36}{81} = 3.556$$

The critical region is $\chi^2 < \chi^2_{0.9}(8) = 3.490$. We accept H_0.

The p-value is 0.105.

3. Suppose a sample $x_1, \ldots, x_{10}$ from a normal distribution $N(\mu_X, \sigma_X^2)$ yielded $\bar{x} = -2.5$ and $s_x^2 = 64$. Suppose a sample $y_1, \ldots, y_{16}$ from a normal distribution $N(\mu_Y, \sigma_Y^2)$ yielded $\bar{y} = 5.5$ and $s_y^2 = 100$. Assume $\alpha = 0.1$.

(a) Test the hypothesis $H_0 : \sigma_X = \sigma_Y$ against $H_1 : \sigma_X \neq \sigma_Y$.

(b) Test the hypothesis $H_0 : \mu_X = \mu_Y$ against $H_1 : \mu_X \neq \mu_Y$. What assumption on σ_X and σ_Y are you making, and why?

(c) Does it appear that the above two distributions are identical? Why or why not?

[Bonus] What can you say about the p-values in the above tests?

Solution:

(a) the test statistic is

$$F = \frac{100}{64} = 1.5625$$

The critical region is $F > F_{0.05}(15, 9) = 3.01$ and $F < 1/F_{0.05}(9, 15) = 1/2.59 = 0.386$. We accept H_0.

The p-value is 0.5051.

(b) We assume $\sigma_X^2 = \sigma_Y^2$, because in part (a) we accepted H_0. Now the test statistic is

$$T = \frac{-2.5 - 5.5}{\sqrt{\frac{9 \cdot 64 + 15 \cdot 100}{10+16-2}} \cdot \sqrt{\frac{1}{10} + \frac{1}{16}}} = -2.1338$$

The critical region is $|T| > t_{0.05}(24) = 1.711$. We accept H_1.

The p-value is 0.0433.

(c) Even though the variances are equal, the means are not equal, so the distributions are not identical.

4. Each of three cars is driven with each of four brands of gasoline. The mileage (miles per gallon) for each car+gasoline combination is recorded in the table below.

	Gasoline			
Car	1	2	3	4
1	27	21	25	23
2	20	17	20	19
3	16	19	15	18

Compute $SS(E)$, $SS(A)$, $SS(B)$ and test the hypotheses

$$H_A: \text{ the choice of car does not affect the gas mileage}$$

and

$$H_B: \text{ the choice of gasoline does not affect the gas mileage}$$

at the level $\alpha = 2.5\%$.

Solution:

$$SS(E) = 30, \qquad SS(A) = 104, \qquad SS(B) = 6$$

Now the test statistics are

$$F_A = \frac{104/2}{30/6} = 10.4, \qquad F_B = \frac{6/3}{30/6} = 0.4$$

The critical regions are $F_A > F_{0.025}(2,6) = 7.26$ and $F_B > F_{0.025}(3,6) = 6.60$. We reject H_A and accept H_B.

5. Let $x_1, \ldots, x_n$ be a random sample from the distribution with probability density function

$$f(x;\theta) = \frac{\theta}{x^2}\, e^{-\frac{\theta}{x}} \qquad \text{for} \quad x > 0$$

where $\theta > 0$ is an unknown parameter.

(a) Find sufficient statistics. (Include as few sufficient statistics as possible.)

(b) Find the Rao-Cramer lower bound on the variance of unbiased estimates of θ.

Solution:

(a) the likelihood function is

$$L(\theta) = \prod_{i=1}^{n} \frac{\theta}{x_i^2}\, e^{-\frac{\theta}{x_i}} = \frac{\theta^n}{\prod_{i=1}^{n} x_i^2}\, e^{-\theta \sum_{i=1}^{n} \frac{1}{x_i}}$$

The (only) test statistic is $\sum_{i=1}^{n} \frac{1}{x_i}$.

(b) The logarithm of the density function is

$$\ln f = \ln\theta - 2\ln x - \frac{\theta}{x}$$

Its first derivative is

$$\frac{\partial}{\partial\theta}\ln f = \frac{1}{\theta} - \frac{1}{x}$$

and its second derivative is

$$\frac{\partial^2}{\partial\theta^2}\ln f = -\frac{1}{\theta^2}$$

The Rao-Cramer lower bound is

$$\mathsf{Var}\hat{\theta} \geq \frac{1}{-n\cdot(-\frac{1}{\theta^2})} = \frac{\theta^2}{n}$$