MA 486-25 (Statistics), Chernov — Midterm test #2

Show your work. — Thu, Apr 27

1. (8 pts) It is claimed that 75% of doctors recommend a certain brand of pain relief medicine. A consumer group doubts this claim (they think the estimate is too high). It conducts a survey that shows that out of 300 doctors, 213 prescribe this brand of medicine.

(a) State the null hypothesis H_0 and the alternative hypothesis H_1.
(b) Find the critical region for the significance level $\alpha = 0.1$.
(c) Which hypothesis do you accept?

Solutions:

$$H_0: \ p = 0.75 \quad H_1: \ p < 0.75$$

$$Z = \frac{\bar{y} - p_0}{\sqrt{p_0(1-p_0)/n}} = -1.6$$

(here $\bar{y} = 213/300 = 0.71$). Critical region:

$$Z < -z_\alpha = -1.282$$

We accept H_1.

2. (12 pts) Let p_m and p_f be the respective proportions of male and female smokers in the US. In a random sample of 800 men, there were 44 smokers, and in a random sample of 600 women, there were 29 smokers. Thest the hypothesis $H_0 : p_m = p_f$ against the alternative $H_1 : p_m \neq p_f$ at 20% level.
(a) Find the critical region.
(b) Which hypothesis do you accept?
(c) [Bonus] By using normal approximation, find an approximate p-value of the test.

Solution:

We have $\bar{y}_1 = 44/800 = 0.055$, $\bar{y}_2 = 29/600 = 0.0483$,

$$\hat{p} = \frac{44 + 29}{800 + 600} = 0.052$$

$$Z = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\hat{p}(1 - \hat{p})(1/n_1 + 1/n_2)}} = 0.55$$

Critical region:

$$|Z| > z_{\alpha/2} = 1.282$$

We accept H_0.

The p-value is $= 2(1 - \Phi(|Z|)) = 2(1 - 0.7088) = 0.5824$

3. (12 pts) Let $x_1, \ldots, x_{64}$ be taken from $N(\mu, 16)$. To test $H_0 : \mu = 20$ against $H_1 : \mu < 20$ let the critical region be $\bar{x} \leq 18.7$.

(a) Find formulas for $\alpha, \beta, K(\mu)$.
(b) Compute α and $K(17)$, $K(18)$, $K(18.5)$, $K(19)$, $K(19.5)$. Sketch the graph of the power function.
(c) Let $\bar{x} = 18.6$. What is the p-value of the test?

Solution:

$$\alpha = \Phi\left(\frac{18.7 - 20}{\sqrt{16}/\sqrt{64}}\right) = \Phi(-2.6) = 0.0047$$

$$\beta = 1 - K(\mu)$$

$$K(\mu) = \Phi\left(\frac{18.7 - \mu}{\sqrt{16}/\sqrt{64}}\right) = \Phi(2(18.7 - \mu))$$

$$K(17) = \Phi(3.4) = 1$$

$$K(18) = \Phi(1.4) = 0.9192$$

$$K(18.5) = \Phi(0.4) = 0.6554$$

$$K(19) = \Phi(-0.6) = 0.2743$$

$$K(19.5) = \Phi(-1.6) = 0.0548$$

The p-value is

$$\Phi(2(18.6 - 20)) = \Phi(-2.8) = 0.0026$$

4. (12 pts) Let X be $N(\mu, 25)$. We need to test the hypethesis $H_0 : \mu_0 = -10$ against $H_1 : \mu_1 = -12$. Let the critical region be $\bar{x} \leq c$. Find values for n and c so that $\alpha = 0.005$ and $K(\mu_1) = 0.999$. Do not forget to find the values of *both* n and c.

Solution:

$$\frac{c - (-10)}{5/\sqrt{n}} = -z_{0.005} = -2.576$$

$$\frac{c - (-12)}{5/\sqrt{n}} = z_{0.001} = 3.090$$

$$n = \frac{(2.576 + 3.090)^2 \cdot 25}{(-12 + 10)^2} = 200.6$$

so $n = 201$, and

$$c = \frac{-10 \cdot 3.090 - 12 \cdot 2.576}{2.576 + 3.090} = -10.909$$

5. (8 pts) Let $x_1, \ldots, x_{19}$ be taken from $N(\mu, \sigma^2)$ to test $H_0 : \mu = 1$ against $H_1 : \mu < 1$. The test results are $\bar{x} = 0.6$ and $s^2 = 0.8$.

(a) Find the critical region.
(b) Do we accept H_0 at the 5% significance level?
(c) What is the approximate p-value of the test?

Solution:

$$T = \frac{0.6 - 1}{\sqrt{0.8}/\sqrt{19}} = -1.95$$

Critical region is

$$T < -t_{0.05}(18) = -1.734$$

We accept H_1.

The p-value is in $(0.025, 0.05)$.

6. (8 pts) Let $x_1, \ldots, x_{16}$ be taken from $N(\mu, \sigma^2)$ to test $H_0 : \sigma = 2$ against $H_1 : \sigma \neq 2$. The test results are $\bar{x} = 4$ and $s^2 = 6.2$.

(a) Find the critical region.
(b) Do we accept H_0 at the 10% significance level?
(c) What is the approximate p-value of the test?

Solution:

$$\chi^2 = \frac{15 \cdot 6.2}{4} = 23.25$$

Critical region is

$$\chi^2 > \chi^2_{0.05}(15) = 25.00 \quad \text{or} \quad \chi^2 < \chi^2_{0.95}(15) = 7.261$$

We accept H_0.

The p-value is in $(0.1, 0.2)$. [The part (c) was graded for extra credit.]

7. (16 pts) Let $x_1, \ldots, x_{11}$ and $y_1, \ldots, y_{13}$ be two independent samples taken from $N(\mu_x, \sigma_x^2)$ and $N(\mu_y, \sigma_y^2)$, respectively. The test results are $\bar{x} = 6$, $s_x^2 = 90$, $\bar{y} = 11$ and $s_y^2 = 100$.

(a) Test the hypothesis $\sigma_x^2 = \sigma_y^2$ against $\sigma_x^2 \neq \sigma_y^2$. Let $\alpha = 5\%$. (Find explicitly the critical region.)
(b) Assume $\sigma_x^2 = \sigma_y^2$, and test the hypothesis $\mu_x = \mu_y$ against $\mu_x \neq \mu_y$. Let $\alpha = 10\%$. (Find explicitly the critical region.)
(c) What is the approximate p-value of the test in part (b)?
(d) Do you conclude that both samples are taken from the same normal distribution?

Solution:

(a) The critical region is

$$s_x^2/s_y^2 = 0.9 > F_{0.025}(10, 12) = 3.37 \quad \text{or} \quad s_y^2/s_x^2 = 1.11 > F_{0.025}(12, 10) = 3.62$$

We accept H_0.

(b) The test statistic is

$$T = \frac{6 - 11}{\sqrt{\frac{10 \cdot 90 + 12 \cdot 100}{22}} \sqrt{\frac{1}{11} + \frac{1}{13}}} = -1.249$$

The critical region is

$$|T| > t_{0.05}(22) = 1.717$$

(c) The p-value is in $(0.2, 0.5)$.

(d) Yes.

8. (12 pts) 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 231, 74, 64, and 31, respectively, for these four categories, would these data support the Mednelian theory? Let $\alpha = 0.05$. What can you say about the p-value of the test?

Solution:

Probabilities of classes are 9/16, 3/16, 3/16, 1/16.

Classes	A	B	C	D
Theory ($= np_i$)	225	75	75	25
Experiment	231	74	64	31

The test statistic is

$$Q = \frac{(231-225)^2}{225} + \frac{(74-75)^2}{75} + \frac{(64-75)^2}{75} + \frac{(31-25)^2}{25} = 3.226$$

The critical region is

$$Q > \chi^2_{0.05}(3) = 7.815$$

We accept H_0.

The p-value is > 0.1.

9. (12 pts) Let X denote the number of fires in a small town on one day. The following 20 observations of X (i.e., the number of recorder fires on 20 consecutive days):

2	0	2	3	2	1	4	1	0	2
4	6	3	0	1	5	0	1	2	1

You need to test the hypothesis that these numbers come from a Poisson distribution.

(a) Estimate the parameter λ by using the maximum likelihood estimator.
(b) Compute the theoretical probabilities for each of the following class intervals: $\{0\}$, $\{1\}$, $\{2\}$, $\{3\}$, $\{4\}$, $\{5, 6, \ldots\}$. Do they satisfy the "minimal safety requirement" $np_i \geq 1$?
(c) With these class intervals, compute the Q test statistic for the chi-square test.
(d) How many degrees of freedom are involved?
(e) Which hypothesis do you accept at the 10% level?

Solution:

$\bar{x} = 40/20 = 2$, so the parameter is $\lambda = \bar{x} = 2$. Next we use the formula

$$p_k = P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

Classes	0	1	2	3	4	$5, 6, \ldots$
Theory ($= np_i$)	2.7	5.4	5.4	3.6	1.8	1.06
Experiment	4	5	5	2	2	2

The test statistic is

$$Q = \frac{(4-2.7)^2}{2.7} + \frac{(5-5.4)^2}{5.4} + \frac{(5-5.4)^2}{5.5} + \frac{(2-3.6)^2}{3.6} + \frac{(2-1.8)^2}{1.8} + \frac{(2-1.06)^2}{1.06} = 2.26$$

The critical region is

$$Q > \chi^2_{0.1}(4) = 7.779$$

We accept H_0.