Problem Set 7 Solution

17.872 TA Jiyoon Kim

Nov. 26, 2003

Bulmer Exercise 9.1

Out of 929 purple × white breed, Mendel observed 705 purple flowers and 224 white flowers. The probability of observed purple flowers .7589 and expected probability is .75. In order to test the hypothesis that the probability of a purple-flowered plant is .75, we first need to get $z-score$ (or d in Bulmer's book).

$$z = \frac{|.7589 - .75|}{\sqrt{.7589(1 - .7589)/929}} = |.63| = \pm.63$$

If you look at the z-table, the probability that z is higher than $\pm$ .63 is .53, therefore, it is not significant.

Bulmer Exercise 9.6

(a) If $P = \frac{1}{2}$, the expected number of successes per throw is 6. What is observed is not the same. First, I calculate the probability of getting the outcomes by dividing each results by 4096. Using $E(x) = \sum xP(x)$, the expected value is obtained, 6.139. Variance of success (σ^2) is $p \times (1-p) \times 12 = 3$. Since we are comparing the sample mean and theoretical mean, we need to divide the variance by the sample size of 4096. Square rooting this result will yield standard deviation. Then,

$$d = \frac{6.139 - 6}{.027} \approx 5.14$$

This is way high d value and statistically significant. Therefore, we can reject the hypothesis.

(b) The way to proceed this question is to calculate the binomial distribution based on a newly observed probability $p = .5116$ for each success. Once it is done, you will have actual results given with the question and expected number of successes based $p = .5116$'s binomial distribution.

You need to use χ^2-distribution of goodness of fit test.

$$\sum \frac{(O - E)^2}{E}$$

1

The result is $\chi^2 = 8.44$ with 11 degrees of freedom. (The book combined the first and the second cells and the eleventh and twelfth cells, since there is no case for the first and the last cell. In that case, it has $\chi^2 = 7$ with 8 degrees of freedom.) This is not significant, either.

S & C 4.8.1
(i) The probability that sample estimate of $\bar{X}$ is within $\pm\$500$,

$$\Pr\left[|\bar{X} - \mu| \leq 500\right] = \Pr\left[\frac{|\bar{X} - \mu|}{\sigma/\sqrt{n}} \leq \frac{500}{\sigma/\sqrt{n}}\right]$$

$$= \Pr\left[|z| \leq \frac{500}{1800/\sqrt{40}}\right] = 1.757$$

The probability that z is less than ± 1.77 is approximately .921.

(ii) Usually, we need 95 percentage probability to ensure the statistical significance, at which z is 1.96. In order to meet this critical value,

$$\Pr\left[\frac{|\bar{X} - \mu|}{\sigma/\sqrt{n}} \leq \frac{500}{1800/\sqrt{n}}\right]$$

$$\frac{500\sqrt{n}}{1800} \geq 1.96$$

$$\sqrt{n} \geq 1.96 \times \tfrac{1800}{500} = 49.787$$

Therefore, n should at least 50 to guarantee the stipulation at 95 percent level.

S & C 4.8.3
Given information of $\mu = 30, \bar{X} = 29.823, \sigma = 10$ and $n = 5110$,
(i) 95 % confidence interval is

$$29.823 \pm 1.96 \times 10/\sqrt{5110} = 29.549, 30.097$$

(ii) Since μ is 30, the confidence interval does cover the μ.

S & C 4.8.8

2

Given information of $\mu = 0$ and $\sigma = .29$.

$$\Pr\left[100|error| \le 5\right] = \Pr\left[\frac{100|error|}{.29/\sqrt{100}} \le \frac{5}{.29/\sqrt{100}}\right]$$

$$= \Pr\left[\frac{|error|}{.29/\sqrt{100}} \le \frac{.5}{.29}\right] = \Pr\left[z \le 1.724\right]$$

The probability that z is less than 1.724 can be looked up in the z-table. We can approximately calculate that $p = .916$ at this z score.

S & C 4.8.9
First, $\sigma = 60$ and $\alpha = .05$. Error is supposed to be within $\pm\$20$.

$$\Pr\left[|\bar{X} - \mu| \le 20\right] = .95$$

$$\Pr\left[\frac{|\bar{X} - \mu|}{60/\sqrt{n}} \le \frac{20}{60/\sqrt{n}}\right] = .95$$

Inside the equation needs to be at least 1.96, therefore,

$$1.96 \le \frac{20}{60/\sqrt{n}}$$

$$\frac{\sqrt{n}}{3} \ge 1.96$$

$$n \ge (1.96 \times 3)^2 = 34.57$$

Therfore, n should be larger than 35. (It would be 36 if you use 2 for simplicity instead 1.96.)

S & C 5.5.1
$\sigma = 20$, $\bar{X} = 56.53$ and $\mu = 50$.
(i) When $n = 25$,

$$z = \frac{\bar{X} - \mu}{\sigma/sqrtn} = \frac{56.53 - 50}{20/\sqrt{25}}$$

$$= 1.6325$$

3

At $z = 1.6325$, $\Pr[z \leq 1.6325] \approx .8968$. Thus, the probability that is getting a larger deviation is 1 - .8968 = .1032.

(ii) When $n = 64$,

$$z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{56.53 - 50}{20/\sqrt{64}}$$

$$= 2.612$$

Now, z is .991 and the probability that is getting larger deviation is 1 - .991 = .009.

S & C 5.5.2
(i) Confidence interval with $n = 25$ is

$$\bar{X} \pm 1.96\sigma/\sqrt{n} = 56.53 \pm 1.96 \times 20/\sqrt{25} = 48.69, 64.37$$

(ii) Confidence interval with $n = 64$ is

$$\bar{X} \pm 1.96\sigma/\sqrt{n} = 56.53 \pm 1.96 \times 20/\sqrt{64} = 51.63, 61.43$$

S & C 5.5.4
$\mu = 500, \sigma = 100, \bar{X} = 472$ and $n = 25$.

$$H_0 : \mu \geq 500$$
$$H_1 : \mu < 500$$

First, calculate z.

$$z = \frac{472 - 500}{100/\sqrt{25}} = -1.4$$

At $z = -1.4$, $p = .5 - .4192 = .0808 \approx .081$.

S & C 5.6.1
(i) Since $\phi = 2$ at $\alpha = .05$, $Z_1 = -2 - 1.96 = -3.96$ and $Z_2 = -2 + 1.96$. $\Pr(Z < Z_1) = 0$ and $\Pr(Z > Z_2) = .5 + .016 = .516$. Therefore, the power is .516.

(ii) With $\phi = 3.5$ and $\alpha = .01$, we only need to change numbers. $Z_1 = -3.5-2.576 = -3.076$ and $Z_2 = -3.5+2.576 = -.924$. $\Pr(Z < Z_1) = 0$ and $\Pr(Z > Z_2) = .5 + .322 = .822$.
(iii) Now, it is one-tailed test. $Z_1 = -3+1.645 = -1.355$ and $\Pr(Z > Z_1) = .5 + .4115 = .9115$.
(iv) $Z_1 = -1.5 + 2.326 = .826$ and $\Pr(Z > Z_1) = .5 - .294 = .206$.

S & C 5.12.1
For 90% confidence interval for χ^2, you need to get $\chi^2_{.95}$ and $\chi^2_{.05}$, under which 90% of probability will be in. In the book, $\chi^2_{.05} = 79.08$ and $\chi^2_{.95)} = 43.19$ with 60 degrees of freedom. (n-1 = 61 - 1 = 60) Assume that $s = 10.2$.

$$\chi^2 = \frac{S^2}{\sigma^2} = \frac{s^2(n-1)}{\sigma^2}$$

$$= \frac{10.2^2 \times 60}{\sigma^2}$$

$$\chi^2_{.95} < \chi^2 < chi^2_{.05} = 43.19 < \chi^2 < 79.08$$

$$= 43.19 < \frac{10.2^2 \times 60}{\sigma^2} < 79.08$$

Then,

$$\sigma^2 > \frac{10.2^2 \times 60}{79.08}$$

$$\sigma^2 < \frac{10.2^2 \times 60}{43.19}$$

These inequalities result in
$\sigma > 8.88$, $\sigma < -8.8$ and $-12.2 < \sigma < 12.2$. The ranges which covers both conditions are $-12.2 < \sigma < -8.8$ and $8.8 < \sigma < 12.2$. Since standard deviation cannot be less than zero, we only take the latter range, $8.8 < \sigma < 12.2$.

S & C 5.12.6
Goodness of fit test. $\sum \frac{(O-E)^2}{E}$ will give you χ^2 with degrees of freedom of 5. $\chi^2 = 6.042$ and with 5 degrees of freedom, we need to have χ^2 at 95% level larger than 11.07. Wince 6.042 is less than 11.07, we fail to reject the null hypothesis.

S & C 7.4.2
The probability of a train arrives on time is .95.
(i) all three arrive on time : $.95^3 = .857$.
(ii) one of the three is late : I think about this in reversed way. I put being late as success and being on time as fail. Using binomial probability, we can calculate $C_{3,1}.05^1 \times .95^2 = .135$.

S & C 7.5.1
(i) Both event 1 and 3 happen : $.11 + .04 = .15$
(ii) Neither even 1 nor 3 happens : $.1 + .4 = .5$
(iii) At least one of events 1 and 2 happens : $1 - .6 = .4$
(iv) At least one of events 1,2 and 3 happens : $1 - .4 = .6$
(v) Verify that in (iv) the probability is less than $p_1 + p_2 + p_3$:
$p_1 = .11 + .03 + .04 + .04 = .22$
$p_2 = .11 + .03 + .08 + .1 = .32$
$p_3 = .11 + .04 + .08 + .2 = .43$
$p_1 + p_2 + p_3 = .97$ And this is larger than .6.

S & C 7.5.2
(i) If p_i is the probability that one event is happening, $q_i = 1 - p_i$ is the probability that an event is not happening. The probability of each of four event happens is 1 minus none of event happens, which will guarantee at least one event happens. Therefore,

$$1 - \sum_{i=1}^4 q_i = 1 - 4q_i \geq .6$$
$$.4 \geq 4q_1 \equiv .1 \geq q_1$$
$$.1 \geq 1 - p_i$$

Therefore, $p_i \geq .9$.
(ii) If events are independent, the probability of all four events happen is $p^4 = .9^4 = .6561$.

S & C 7.6.3
$n = 96$ and $p = .4$.
(i) Standard deviation of the number is $\sigma = \sqrt{.4 \times .6 \times 96} = 4.8$
(ii) Standard deviation of the percentage of successes is $\sigma = \sqrt{\frac{.4 \times .6}{96}} = .05$.
Therefore, 5 %.

S % C 7.6.4

Since $p = .1$ and $\sigma = .01$, we can set up the equation that

$$.01 = \sqrt{\frac{.1 \times .9}{n}}$$

Therefore, $n = 900$.

S % C 7.6.5

For $n = 10, p = .5$,

(i) Using binomial distribution, you count from 0 to 4 successes.

$$C_{10,0}.5^0.5^{10} + C_{10,1}.5^1.5^9 + C_{10,2}.5^2.5^8 + C_{10,3}.5^3.5^7 + C_{10,4}.5^4.5^8 = .377$$

(ii) the normal approximation corrected for continuity is

$$z_c = (|r - np| - .5)/\sqrt{npq} = (|4 - 5| - .5)/1.581 = .5/1.581 = .316$$

$\Pr(z_c > .316)$ is same as the probability of $\Pr(z_c < -.316)$ and it is $.5 - \Pr(z_c < .316) \approx .5 - .124 = .376$.

(iii) the normal approximation uncorrected for continuity is

$$z = (|r - np|)/\sqrt{npq} = (|4 - 5|)/1.581 = 1/1.581 = .6325$$

$\Pr(z > .6325) = .5 - .237 = .263$.