EXERCISES IN STATISTICS

Series A, No. 10

1. A horticulturist considers that a batch of seeds is worth sowing if 50% of the resulting flowers are going to be pure white. To test the worth of a particular batch, he sows eight seeds with the intention of sowing the remainder if at least four of the eight plants have white flowers. Find the probability of his making a wrong decision (a) if 25% of the seeds are of the white variety, and (b) if 75% of the seeds are of the white variety.

Answer Let p be the probability that a seed selected at random will produced a white flower, and let x be the number of white flowers in the test sample of $n = 8$ seeds. We may assume that

$$x \sim b(n, p) = \frac{n!}{(n-x)!x!} p^x (1-p)^{n-x}.$$

(a) Let $p = 1/4$. The wrong decision, which is to sow the seeds, is made if $x \in \{4, 5, 6, 7, 8\}$. Observe that $P(x \in \{4, 5, 6, 7, 8\}) = 1 - P(x \in \{0, 1, 2, 3, \})$. Then

$$\begin{aligned} P(x \in \{0,1,2,3,\}) &= \sum_{x=0}^{4} \frac{n!}{(n-x)!x!} p^x (1-p)^{n-x} \\ &= \left(\frac{3}{4}\right)^8 + 8\left(\frac{1}{4}\right)\left(\frac{3}{4}\right)^7 + \frac{8.7}{2.1}\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^6 + \frac{8.7.6}{3.2.1}\left(\frac{1}{4}\right)^3\left(\frac{3}{4}\right)^5 \\ &= \left(\frac{1}{4}\right)^8 \left\{3^8 + 8.3^7 + 4.7.3^6 + 8.7.3^5\right\}. \end{aligned}$$

We have

$$3^5 = 243, \quad 3^6 = 729, \quad 3^7 = 2187, 3^8 = 6561, \quad \text{and} \quad 4^8 = 65536.$$

Therefore

$$\begin{aligned} P(x \in \{0,1,2,3,\}) &= \frac{1}{65536}\left\{6561 + (8 \times 2187) + (28 \times 729) + (56 \times 243)\right\} \\ &= 0.8877, \end{aligned}$$

and

$$P(x \in \{4,5,6,7,8\}) = 1 - 0.8877 = 0.1153$$

(b) Let $p = 3/4$. The wrong decision, which is not to sow the seeds, is made if $x \in \{0, 1, 2, 3\}$. Then

$$\begin{aligned}&P(x \in \{0,1,2,3,\})\\&= \left(\frac{1}{4}\right)^8 + 8\left(\frac{1}{4}\right)^7\left(\frac{3}{4}\right) + \frac{8.7}{2.1}\left(\frac{1}{4}\right)^6\left(\frac{3}{4}\right)^2 + \frac{8.7.6}{3.2.1}\left(\frac{1}{4}\right)^5\left(\frac{3}{4}\right)^3\\&= \left(\frac{1}{4}\right)^8\Big\{1 + (8.3) + (4\times 7\times 9) + (8\times 7\times 27)\Big\} = \frac{1789}{65536} = 0.0273.\end{aligned}$$

(c) If we were to assert the null hypothesis $H_0 : p = 0.5$, then the probability of a Type II error would be

$$\begin{aligned}&b(0) + b(1) + b(2) + b(3) = \left(\frac{1}{2}\right)^8\left\{1 + 8 + \frac{8\times 7}{2\times 1} + \frac{8\times 7\times 6}{3\times 2\times 1}\right\}\\&= \frac{1}{256}\Big\{1 + 8 + 28 + 56\Big\} = \frac{93}{256} = 0.3643.\end{aligned}$$

2. In the previous election, 65% of the poll voted for the Progressive Party. In a recent opinion survey, 280 out of 400 people have said that they intend to vote for the Progressive Party. Is there any substantial evidence to suggest that the party has increased its support? Hint: use the fact that

$$\left(\frac{x}{n} - p\right)\Bigg/\sqrt{\frac{pq}{n}}$$

is distributed approximately as a standard normal variate.

Answer If $x \sim b(n, p)$ then $E(x) = np$, $V(x) = npq$ which implies that

$$E\left(\frac{x}{n}\right) = p \qquad \text{and} \qquad V\left(\frac{x}{n}\right) = \frac{1}{n^2}V(x) = \frac{pq}{n}.$$

Also, by the central limit theorem,

$$\lim(n \to \infty)\left(\frac{x}{n} - p\right)\Bigg/\sqrt{\frac{pq}{n}} = z \sim N(0, 1).$$

Under the null hypothesis, it is presumed that

$$E\left(\frac{x}{n}\right) = p = 0.65.$$

From the sample

$$\frac{x}{n} = \frac{280}{400} = 0.7.$$

Therefore

$$\left(\frac{x}{n} - p\right) \Big/ \sqrt{\frac{pq}{n}} = \left(0.65 - 0.70\right) \Big/ \sqrt{\frac{0.65 \times 0.35}{400}}$$
$$= -\frac{0.05 \times 100}{\sqrt{5.6875}} = 2.096.$$

On the basis of a one-tailed test using the the $N(0, 1)$ distribution, this is significant 5% level.

3. Imagine that a random sample of size 160 is drawn from a normal population where the standard deviation is 25 in order to test the null hypothesis that the mean is 104 against the alternative that it is 100. Calculate the probability of a Type II error in a one-tailed test with a 5% significance level. What sample size is required for the probability of the Type II error to be 0.05?

Answer We assume that $x \sim N(\mu, \sigma^2)$. Given the null hypothesis $H_0 : \mu = \mu_0$ and the alternative hypothesis $H_1 : \mu = \mu_1 < \mu_0$, we adopt the following decision rule:

$$\begin{array}{lll} d_1 & \text{if} & \dfrac{\bar{x} - \mu_0}{\sigma\sqrt{n}} \leq -b \\ d_0 & \text{if} & \dfrac{\bar{x} - \mu_0}{\sigma\sqrt{n}} > -b \end{array}$$

The null hypothesis and the alternative hypothesis are respectively

$$\begin{array}{ll} H_0 : \mu_0 = 104, & \sigma = 25, \\ H_1 : \mu_0 = 100, & \sigma = 25. \end{array}$$

The critical value for a one-tailed test at the 5% level is the value of b, obtained from the $N(z; 0, 1)$ tables which satisfies

$$P(z > -b) = 0.95 \Longrightarrow -b = -1.65$$

Now

$$\frac{\bar{x} - \mu_0}{\sigma\sqrt{n}} > -b \Longleftrightarrow \bar{x} - \mu_0 > -b\frac{\sigma}{\sqrt{n}}$$
$$\Longleftrightarrow \bar{x} - \mu_1 > -b\frac{\sigma}{\sqrt{n}} + (\mu_0 - \mu_1)$$
$$\Longleftrightarrow \frac{\bar{x} - \mu_1}{\sigma\sqrt{n}} > -b + \frac{\mu_0 - \mu_1}{\sigma/\sqrt{n}}.$$

Putting the the number in place gives

$$-b + \frac{\mu_0 - \mu_1}{\sigma/\sqrt{n}} = -1.65 + \frac{104 - 100}{25/sqrt160} = 0.373.$$

Under the alternative hypothesis H_1, we have

$$\frac{\bar{x} - \mu_1}{\sigma\sqrt{n}} = z \sim N(0, 1)$$

and so the probability of a Type II error is

$$P(d_0|\mu = \mu_1) = P(z > 0.373) = 1 - (0.5 + 0.1443) = 0.3557$$

nest we are asked to find a value for the sample size n such that under h_1 we should have

$$P\left(z = \frac{\bar{x} - \mu_1}{\sigma\sqrt{n}} > \lambda = -b + \frac{\mu_0 - \mu_1}{\sigma\sqrt{n}}\right) = 0.05.$$

Now

$$P(z > \lambda) = 0.05 \Longrightarrow \lambda = 1.65$$

so we must solve the equation

$$\lambda = -b + \frac{\mu_0 - \mu_1}{\sigma\sqrt{n}}$$

for the unknown n. We have

$$\sqrt{n} = \frac{(\lambda + b)\sigma}{\mu_0 - \mu_1} = \frac{(1.65 + 1.65) \times 25}{4} = 20.625 \Longrightarrow n = 426.$$