Exercise 3 Solutions

1. Let $X$ denote the number having the disease in the sample, $X \sim B(n, p)$.

Test $H_0 : p = 0.1$ against the alternative $H_1 : p > 0.1$. at significance level 0.05.

For large $n$, $X$ is approximately $N(0.1n, 0.09n)$. The approximate rejection region is

$$X > 0.1n + 1.645\sqrt{0.09n}.$$ 

If $p = 0.13$ then $X$ is approximately $N(0.13n, 0.1131n)$. We want

$$0.95 \leq P(X > 0.1n + 1.645\sqrt{0.09n} | p = 0.13)$$

Thus

$$-1.645 \geq (0.1n + 1.645\sqrt{0.09n} - 0.13n)/\sqrt{0.1131n}$$

$$n \geq (34.8906)^2 = 1217.36$$

A sample of at least 1218 is needed.

2. $X \sim B(n, p)$. we want to test $H_0 : p = p_0$ against $H_1 : p = p_1 (> p_0)$. The rejection region according to the Neyman-Pearson Lemma is

$$\frac{\binom{n}{x} p_0^x (1-p_0)^{n-x}}{\binom{n}{x} p_1^x (1-p_1)^{n-x}} < c$$

for some constant $c$. This is equivalent to

$$x \log \left( \frac{p_0(1-p_1)}{p_1(1-p_0)} \right) < k$$

Since $\log \left( \frac{p_0(1-p_1)}{p_1(1-p_0)} \right) < 0$ the rejection region is

$$x > k_1,$$

for some constant $k_1$. The (approximate) rejection region, using the normal approximation is

$$X > np_0 + z_{1-\alpha} \sqrt{np_0(1-p_0)}.$$
3. Let $X$ be the breaking point of the string in standard technique and $Y$ be the breaking point of the string in new technique. Because the sample size is small, to perform a two-sample $t$-test, we have to assume

$$X \sim N(\mu_X, \sigma^2_X) \quad \text{and} \quad Y \sim N(\mu_Y, \sigma^2_Y).$$

Note that $\mu_X$ and $\mu_Y$ are the mean breaking points of the string in the standard technique and the new technique respectively.

(a) Test $H_0 : \sigma^2_X = \sigma^2_Y$ against the general alternative.

We have

i. $\bar{x} = 138$, \quad $\bar{y} = 143$;

ii. $s^2_x = 133.5$, \quad $s^2_y = 68.5$;

iii. $s^2 = (4 \times s^2_x + 4 \times s^2_y)/8 = 101$,

The test statistic is $f = 133.5/68.5 = 1.949$. The $p$-value is

$$p = 2 \times P(F_{4,4} \geq 1.949) = 0.5340.$$

The data are consistent with the two populations having common variance. This test requires the normal assumption for the shape of the populations.

(b) The hypothesis to be tested is

$$H_0 : \mu_X = \mu_Y \quad \text{vs} \quad H_1 : \mu_X < \mu_Y.$$

$$t_{m,n} = (\bar{x} - \bar{y})/ \left( s \sqrt{1/5 + 1/5} \right) = -0.786646.$$

The $p$-value is given by

$$p = P(t_8 \leq -0.786646) = 0.22708.$$

**Conclusion:** The data are consistent with the null hypothesis, that is, there is no evidence to support the contention of the research department.