Exercise 8.28

In a study of the relationship between birth order and college success, an investigator found that 126 in a sample of 180 college graduates were firstborn or only children; in a sample of 100 nongraduates of comparable age and socioeconomic background, the number of firstborn or only children was 54. Estimate the difference in the proportions of firstborn or only children for the two populations from which these samples were drawn. Give a bound for the error of estimation.

Exercise 8.44

Let Y have probability density function

$$f_Y(y) = \begin{cases} \dfrac{2(\theta - y)}{\theta^2}, & 0 < y < \theta \\ 0, & \text{elsewhere} \end{cases}$$

a Show that Y has distribution function

$$F_Y(y) = \begin{cases} 0, & y \leq 0 \\ \dfrac{2y}{\theta} - \dfrac{y^2}{\theta^2}, & 0 < y < \theta \\ 1, & y \geq \theta \end{cases}$$

b Show that Y/θ is a pivotal quantity.

c Use the pivotal quantity from part (b) to find a 90% lower confidence limit for θ.

Exercise 8.47

Assume that $Y_1, Y_2, \ldots, Y_n$ is a sample of size n from an exponential distribution with mean θ.

- **a** Use the method of moment-generating functions to show that $2\sum_{i=1}^{n} Y_i/\theta$ is a pivotal quantity and has a χ^2 distribution with $2n$df.
- **b** Use the pivotal quantity $2\sum_{i=1}^{n} Y_i/\theta$ to derive a 95% confidence interval for θ.
- **c** If a sample of size $n = 7$ yields $\bar{y} = 4.77$, use the result from part (b) to give a 95% confidence interval for θ.

Exercise 8.65

For a comparison of the rates of defectives produced by two assembly lines, independent random samples of 100 items were selected from each line. Line A yielded 18 defectives in the sample, and line B yielded 12 defectives.

- **a** Find a 98% confidence interval for the true difference in proportions of defectives for the two lines.
- **b** Is there evidence here to suggest that one line produces a higher proportion of defectives than the other?

1 Confidence Intervals for Common Distributions

1.1 Normal Distribution

1.1.1 Case 1: Known Variance σ^2

Suppose $X_1, X_2, \ldots, X_n$ are independent and identically distributed (i.i.d.) random variables from $N(\mu, \sigma^2)$, where σ^2 is known.

Pivotal Quantity The standardized sample mean:

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

Constructing the Confidence Interval For a confidence level $1-\alpha$, find $z_{\alpha/2}$ such that $P(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}) = 1-\alpha$.

Then,

$$P\left(-z_{\alpha/2} \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq z_{\alpha/2}\right) = 1-\alpha$$

Solving for μ:

$$P\left(\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

Thus, the $(1-\alpha)100\%$ confidence interval for μ is:

$$\left(\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \quad \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

1.1.2 Case 2: Unknown Variance σ^2

When σ^2 is unknown, we estimate it using the sample variance S^2.

Pivotal Quantity The t-statistic:

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$$

Constructing the Confidence Interval Find $t_{\alpha/2,n-1}$ such that $P(-t_{\alpha/2,n-1} \leq T \leq t_{\alpha/2,n-1}) = 1-\alpha$.

Then,

$$P\left(-t_{\alpha/2,n-1} \leq \frac{\bar{X} - \mu}{S/\sqrt{n}} \leq t_{\alpha/2,n-1}\right) = 1-\alpha$$

Solving for μ:

$$P\left(\bar{X} - t_{\alpha/2,n-1}\frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}\right) = 1-\alpha$$

Thus, the $(1-\alpha)100\%$ confidence interval for μ is:

$$\left(\bar{X} - t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}, \quad \bar{X} + t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}\right)$$

1.1.3 Confidence Interval for Variance σ^2

Pivotal Quantity The scaled sample variance:

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

Constructing the Confidence Interval Find $\chi^2_{\alpha/2,n-1}$ and $\chi^2_{1-\alpha/2,n-1}$ such that:

$$P\left(\chi^2_{1-\alpha/2,n-1} \leq \chi^2 \leq \chi^2_{\alpha/2,n-1}\right) = 1-\alpha$$

Solving for σ^2:

$$P\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}}\right) = 1-\alpha$$

Thus, the $(1-\alpha)100\%$ confidence interval for σ^2 is:

$$\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}}, \quad \frac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}}\right)$$

1.2 Exponential Distribution

Suppose $X_1, X_2, \ldots, X_n$ are i.i.d. from an exponential distribution with parameter λ (rate parameter).

Pivotal Quantity The sum of exponential random variables follows a gamma distribution:

$$\frac{2\sum_{i=1}^{n} X_i}{\lambda} = \frac{2n\bar{X}}{\lambda} \sim \chi^2_{2n}$$

Constructing the Confidence Interval Find $\chi^2_{\alpha/2,2n}$ and $\chi^2_{1-\alpha/2,2n}$ such that:

$$P\left(\chi^2_{\alpha/2,2n} \leq \frac{2n\bar{X}}{\lambda} \leq \chi^2_{1-\alpha/2,2n}\right) = 1-\alpha$$

Solving for λ:

$$P\left(\frac{2n\bar{X}}{\chi^2_{1-\alpha/2,2n}} \leq \lambda \leq \frac{2n\bar{X}}{\chi^2_{\alpha/2,2n}}\right) = 1-\alpha$$

Thus, the $(1-\alpha)$ confidence interval for λ is:

$$\left(\frac{2n\bar{X}}{\chi^2_{1-\alpha/2,2n}}, \quad \frac{2n\bar{X}}{\chi^2_{\alpha/2,2n}}\right)$$

Exercise 8.28

The point estimate is given by the difference of the sample proportions: $.70 - .54 = .16$ and an error bound is $2\sqrt{\frac{.7(.3)}{180} + \frac{.54(.46)}{100}} = .121$.

Exercise 8.44

a $F_Y(y) = P(Y \le y) = \int_0^y \frac{2(\theta-t)}{\theta^2} dt = \frac{2y}{\theta} - \frac{y^2}{\theta^2}, 0 < y < \theta.$

b The distribution of $U = Y/\theta$ is given by $F_U(u) = P(U \le u) = P(Y \le \theta u) = F_Y(\theta u) = 2u - u^2, 0 < u < 1$. Since this distribution does not depend on $\theta, U = Y/\theta$ is a pivotal quantity.

c Set $P(U \le a) = F_Y(a) = 2a - a^2 = .9$ so that the quadratic expression is solved at $a = 1 - \sqrt{.10} = .6838$ and then the 90% lower bound for θ is $Y/.6838$.

Exercise 8.47

Note that for all i, the mgf for Y_i is $m_Y(t) = (1-\theta t)^{-1}, t < 1/\theta$.

a Let $U = 2\sum_{i=1}^n Y_i/\theta$. The mgf for U is

$$m_U(t) = E\left(e^{tU}\right) = [m_Y(2t/\theta)]^n = (1-2t)^{-n}, t < 1/2$$

This is the mgf for the chi-square distribution with $2n$ degrees of freedom. Thus, U has this distribution, and since the distribution does not depend on θ, U is a pivotal quantity.

b Similar to part b in Ex. 8.46 , let $\chi^2_{.975}, \chi^2_{.025}$ be percentage points from the chi-square distribution with $2n$ degrees of freedom such that

$$P\left(\chi^2_{.975} \le 2\sum_{i=1}^n Y_i/\theta \le \chi^2_{.025}\right) = .95.$$

So

$$\left(\frac{2\sum_{i=1}^n Y_i}{\chi^2_{.975}}, \frac{2\sum_{i=1}^n Y_i}{\chi^2_{.025}}\right) \text{ represents a 95\%CI for } \theta.$$

c The CI is $\left(\frac{2(7)(4.77)}{26.1190}, \frac{2(7)(4.77)}{5.62872}\right)$ or $(2.557, 11.864)$.

Exercise 8.65

a The 98%CI is, with $z_{.01} = 2.326$, is

$$(.18 - .12) \pm 2.326\sqrt{\frac{18(.82) + 12(.88)}{100}} \text{ or } .06 \pm .117 \text{ or } (-.057, .177).$$

b Since the interval contains both positive and negative values, it is likely that the two assembly lines produce the same proportion of defectives.