Exercise

Prove that a t-distributed random variable T_n with n degrees of freedom converges in distribution to a standard normal random variable Z as $n \to \infty$.

Exercise 10.42

A random sample of 500 measurements on the length of stay in hospitals had sample mean 5.4 days and sample standard deviation 3.1 days. A federal regulatory agency hypothesizes that the average length of stay is in excess of 5 days.

a Do the data support this hypothesis? Use $\alpha = .05$.

b How large should the sample size be if we require that $\alpha = .01$ and $\beta = .05$ when $\mu_a = 5.5$?

Exercise 10.97

Let $Y_1, Y_2, \ldots, Y_n$ be independent and identically distributed random variables with discrete probability function given by

	y		
	1	2	3
$p(y \mid \theta)$	θ^2	$2\theta(1-\theta)$	$(1-\theta)^2$

where $0 < \theta < 1$. Let N_i denote the number of observations equal to i for $i = 1, 2, 3$.

a Derive the likelihood function $L(\theta)$ as a function of N_1, N_2, and N_3.

b Find the most powerful test for testing $H_0 : \theta = \theta_0$ versus $H_a : \theta = \theta_a$, where $\theta_a > \theta_0$. Show that your test specifies that H_0 be rejected for certain values of $2N_1 + N_2$.

c How do you determine the value of k so that the test has nominal level α? You need not do the actual computation. A clear description of how to determine k is adequate.

d Is the test derived in parts (**a**)-(**c**) uniformly most powerful for testing $H_0 : \theta = \theta_0$ versus $H_a : \theta > \theta_0$? Why or why not?

Exercise 10.42

a Let $\mu =$ mean length of stay in hospitals. Then, for $H_0 : \mu = 5, H_a : \mu > 5$, the large sample test statistic is $z = 2.89$. With $\alpha = .05$, $z_{.05} = 1.645$ so we reject H_0.

b Using the sample size formula, we have

$$n \geq \frac{(z_\alpha + z_\beta)^2 \sigma^2}{(\mu_a - \mu_0)^2} = 607.37,$$

so a sample size of 608 will provide the desired levels.

Exercise 10.97

Note that (N_1, N_2, N_3) is trinomial (multinomial with $k = 3$) with cell probabilities as given in the table.

a The likelihood function is simply the probability mass function for the trinomial:

$$L(\theta) = \frac{n!}{n_1!n_2!n_3!}\theta^{2n_1}[2\theta(1-\theta)]^{n_2}(1-\theta)^{2n_3}, 0 < \theta < 1, n = n_1 + n_2 + n_3.$$

b Using part a, the best test for testing $H_0 : \theta = \theta_0$ vs. $H_a : \theta = \theta_a, \theta_0 < \theta_a$, is

$$\frac{L(\theta_0)}{L(\theta_a)} = \left(\frac{\theta_0}{\theta_a}\right)^{2n_1+n_2}\left(\frac{1-\theta_0}{1-\theta_a}\right)^{n_2+2n_3} < k$$

Since we have that $n_2 + 2n_3 = 2n - (2n_1 + n_2)$, the RR can be specified for certain values of $S = 2N_1 + N_2$. Specifically, the log-likelihood ratio is

$$S \ln\left(\frac{\theta_0}{\theta_a}\right) + (2n - S)\ln\left(\frac{1-\theta_0}{1-\theta_a}\right) < \ln k,$$

or equivalently

$$S > \left[\ln k - 2n \ln\left(\frac{1-\theta_0}{1-\theta_a}\right)\right] \times \left[\ln\left(\frac{\theta_0(1-\theta_a)}{\theta_a(1-\theta_0)}\right)\right]^{-1} = c.$$

So, the rejection region is given by $\{S = 2N_1 + N_2 > c\}$.

c To find a size α rejection region, the distribution of (N_1, N_2, N_3) is specified and with $S = 2N_1 + N_2$, a null distribution for S can be found and a critical value specified such that $P(S \geq c \mid \theta_0) = \alpha$. Observe that the null distribution of the statistic S is equivalent to a binomial variable with $2n$ trials and probability θ_0.

d Since the RR doesn't depend on the value of the alternative θ_a, it is a UMP test.