2.4 Computational Hypothesis Testing

This project requires an understanding of the Part IB Statistics course.

In the standard framework for hypothesis testing, we compare the observed value of a test statistic to what we would expect it to be under the null hypothesis. For some tests, like the t-test or the χ^2 test, we can look up statistical tables to help us perform this comparison. For other tests, tables are not available; in this project we will see how to use simulation instead of tables.

Student's *t*-test

To illustrate the idea, consider the t-test. Let $X_1, \ldots, X_m$ be independent random variables, all $\mathrm{N}(\mu_1, \sigma^2)$, and let $Y_1, \ldots, Y_n$ be independent random variables, all $\mathrm{N}(\mu_2, \sigma^2)$. Suppose σ^2 is unknown. We wish to test

$$H_0 : \; \mu_1 = \mu_2 \quad \text{against} \quad H_1 : \; \mu_1 \neq \mu_2.$$

The standard method for testing these hypotheses can be expressed as a likelihood ratio test, which says "Reject H_0 if the likelihood ratio $\mathrm{LR}(H_1, H_0)$ is greater than C", for some suitable value of C.

Question 1 Calculate $\mathrm{LR}(H_1, H_0)$. Show that the likelihood ratio test can be expressed as "Reject H_0 if $|T| > D$", for some suitable value of D, where

$$T = \frac{(\bar{X} - \bar{Y})(m^{-1} + n^{-1})^{-1/2}}{\sqrt{S_{XX} + S_{YY}}(m + n - 2)^{-1/2}}, \quad \bar{X} = \frac{1}{m}\sum_{i=1}^{m} X_i, \quad S_{XX} = \sum_{i=1}^{m}(X_i - \bar{X})^2$$

and $\bar{Y}$ and S_{YY} are defined similarly.

Recall that $\bar{X} \sim \mathrm{N}(\mu_1, m^{-1}\sigma^2)$ and $S_{XX} \sim \sigma^2\chi^2_{m-1}$, and that $\bar{Y}$ and S_{YY} have similar distributions, and that these four random variables are independent.

The *size* of the test is the probability of a Type I error, i.e., the probability that we reject H_0 when it is true. Under the standard framework for hypothesis testing, we choose C (or equivalently D) so that the test has size α, typically $\alpha = 5\%$. The *power* of the test, given values for $\mu_1 \neq \mu_2$ and σ^2, is the probability that a Type II error does not occur, i.e., the probability that we reject H_0 given these parameter values.

Question 2 Show that under H_0 the distribution of T does not depend on μ_1, μ_2 or σ^2. Explain how one can use this fact to choose D so that the test has a given size α. Use tables to find the critical value D such that for $m = 15$ and $n = 20$ the test has size $\alpha = 5\%$.

For this test, the critical value D (and also the power of the test) can be looked up in standard tables. If we do not have tables accessible, or if the distribution of the test statistic is not standard, we can use computer simulation to draw samples from the distribution of T.

Question 3 Estimate the value D such that $\mathbb{P}(|T| > D) = 5\%$ by generating samples of 10,000 independent realisations of T. Compare to your answer to Question 2.

Programming hint: If your programming environment does not support the generation of normal or χ^2 random variables, see the appendix for notes on how to generate them.

Inhomogenous variances

Suppose that $(X_1, \ldots, X_m)$ is a sample of independent $\mathrm{N}(\mu_1, \sigma_1^2)$ random variables, and that $(Y_1, \ldots, Y_n)$ is a sample of independent $\mathrm{N}(\mu_2, \sigma_2^2)$ random variables. Suppose all of the parameters are unknown, and we wish to test

$$H_0 : \mu_1 = \mu_2 \text{ and } \sigma_1^2 = \sigma_2^2 \quad \text{against} \quad H_1 : \mu_1 \neq \mu_2 \text{ or } \sigma_1^2 \neq \sigma_2^2.$$

Question 4 Calculate $\mathrm{LR}(H_1, H_0)$. Show that the likelihood ratio test can be expressed as "Reject H_0 if $|T| > D$", where T is a function (which you should find) of $\bar{X}$, $\bar{Y}$, S_{XX} and S_{YY}.

Question 5 Show that under H_0 the distribution of T does not depend on μ_1, μ_2, σ_1^2 or σ_2^2. Find the critical value D for a test of size $\alpha = 5\%$ when $m = 15$ and $n = 20$.

We can also use simulation to find the power of the test. Suppose H_1 is true, and the parameters μ_1, μ_2, σ_1^2 and σ_2^2 are given. By simulating a random sample of T given these parameters, we can estimate the power of the test, which is just $\mathbb{P}(T > D | H_1)$.

Question 6 Find the power of the likelihood ratio test under the alternative $\mu_1 = 0$, $\mu_2 = 0$, $\sigma_1 = 1$, $\sigma_2 = 3$.

Question 7 With the same μ_1, μ_2, σ_1^2 and σ_2^2, and $\alpha = 5\%$, find the power of the likelihood ratio test for a range of values of m and n. Subject to $m + n = 10$, what values of m and n yield the most powerful test?

Theory says that the distribution of $2 \log \mathrm{LR}(H_1, H_0)$ under H_0 should approach that of a χ^2 random variable with two degrees of freedom. A standard way to compare two distributions visually is with the quantile-quantile plot: given cumulative distribution functions F and G, one plots for each $p \in [0, 1]$ the point $(F^{-1}(p), G^{-1}(p))$. If one does not know F exactly, one can approximate it by the empirical cumulative distribution function $\hat{F}$ of a random sample $(Z_1, \ldots, Z_N)$ drawn from F:

$$\hat{F}(z) = \frac{1}{N} \sum_{i=1}^{N} 1[Z_i \leqslant z]$$

This has the convenient property that

$$\hat{F}^{-1}(i/N) = Z_{(i)}$$

where $Z_{(i)}$ is the ith order statistic, i.e., the ith smallest value in the sample.

Question 8 Compare the distribution of $2 \log \mathrm{LR}(H_1, H_0)$ to that of a χ^2 random variable with two degrees of freedom by means of a quantile-quantile plot.

Notes on generating random variables

Let U and V be independent uniform random variables on $[0,1]$. The MATLAB routine **rand** generates such random variables.

Let $X = -\lambda^{-1}\log U$. Then X is an exponential random variable with mean λ^{-1}. If $E_1,\ldots,E_n$ are independent exponential random variables with mean λ^{-1}, then $E_1+\cdots+E_n$ is a Gamma random variable $\Gamma(n,\lambda)$.

Let $\theta = 2\pi U$ and $R = -2\log V$. Then $\sqrt{R}\cos\theta$ and $\sqrt{R}\sin\theta$ are independent standard normal random variables (i.e., with mean 0 and variance 1). If $A_1,\ldots,A_d$ are independent standard normal random variables, then $A_1^2+\cdots+A_d^2$ is a χ_d^2 random variable (with d degrees of freedom). Note that $\chi_d^2 \sim \Gamma(d/2, 1/2)$.

