STA 114: Statistics

Lab 8

Consider data $X = (X_1, \cdots, X_n)$ and $Y = (Y_1, \cdots, Y_m)$ modeled as $X_i \overset{\text{IID}}{\sim} \mathsf{Normal}(\mu_1, \sigma^2)$ and $Y_j \overset{\text{IID}}{\sim} \mathsf{Normal}(\mu_2, \sigma^2)$, where the quantity of interest is $\eta = \mu_1 - \mu_2$. We know that a $100(1-\alpha)\%$ ML confidence interval for η is $\bar{x} - \bar{y} \mp z_{n+m-2}(\alpha)\sqrt{(\frac{1}{n} + \frac{1}{m})\frac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2}}$. Therefore a size α ML test for $H_0 : \eta = 0$ against $H_1 : \eta \neq 0$ is of the form:

$$\textit{reject } H_0 \text{ if and only if } \frac{|\bar{x} - \bar{y}|}{\sqrt{(\frac{1}{n} + \frac{1}{m})\frac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2}}} > z_{n+m-2}(\alpha),$$

which is the so called *equal-variance t-test.*

Our model on X, Y is really a special case of $X_i \overset{\text{IID}}{\sim} g(x_i - \eta)$, $Y_j \overset{\text{IID}}{\sim} g(y_j)$ where $\eta \in (-\infty, \infty)$ and g is some pdf on $(-\infty, \infty)$. The normal model described above obtains when we take g to be $\mathsf{Normal}(\mu_2, \sigma^2)$. But for most scientific applications, it's hard to ascertain that g is a normal density. So we should possibly consider as parameter space

$$\Theta = \{(\eta, g) : -\infty < \eta < \infty, g \text{ some pdf on } (-\infty, -\infty)\},$$

and our null hypothesis is $H_0 : \theta \in \Theta_0 = \{(0, g) : g \text{ some pdf on } (-\infty, \infty)\}$ and $H_1 : \theta \in \Theta_1 = \Theta \setminus \Theta_0$. Does a equal-variance t-test retain its size for this testing problem, where we do not assume normality of g? We'll investigate this with $\alpha = 0.05$ by using Monte Carlo simulations.

Simulations for $g = \mathsf{Normal}(0, 1)$ and $g = t(1)$

We'd first use simulations to see whether we match the theoretical calculations for the normal model.

> Task 1. Set $n = 10$, $m = 15$, $\mu_1 = \mu_2 = 0$ and $\sigma = 1$. Simulate draws for x and y: `x <- rnorm(10); y <- rnorm(15)`. Check whether the t-test (with $\alpha = 0.05$) rejects $H_0 : \eta = 0$. Now repeat this 10,000 times and get the proportion of times H_0 was rejected. This should roughly equal the size 5% of the test.

> Task 2. Repeat the above task but now simulate data as: `x <- rt(10, 1); y <- rt(15, 1)`.

A test with the correct size

Consider the statistic W which counts the number of pairs (X_i, Y_j) for which $Y_j \leq X_i$. Let w denote the value of W recorded from data. Then w ranges between 0 (all x_i's are smaller than every y_j's) and nm (every x_i is larger than all y_j's). When $\eta = 0$, the distribution of W does not depend on the pdf g! This is because when $\eta = 0$, the X_i's and Y_j's are IID from the pdf g, and therefore, all arrangements of X_i's and Y_j's are equally likely. Since the distribution of W only depends on the arrangements of X_i's and Y_j's, it is unaffected by the exact form of g.

To see this let $n = 2$ and $m = 3$. Then the possible arrangements of X_i's and Y_j's in data can be described by the $\binom{5}{2} = 10$ strings shown on the first column of the following table. The string `xyxyy` indicates that when put in ascending order, the first and the third entries are from X_i's, whereas the second, fourth and fifth entries are from the Y_j's. The second column of the table shows the corresponding values of W. Each arrangement has $\frac{1}{10}$ probability.

Arrangements	W	Probability
`xxyyy`	0	$\frac{1}{10}$
`xyxyy`	1	$\frac{1}{10}$
`xyyxy`	2	$\frac{1}{10}$
`xyyyx`	3	$\frac{1}{10}$
`yxxyy`	2	$\frac{1}{10}$
`yxyxy`	3	$\frac{1}{10}$
`yxyyx`	4	$\frac{1}{10}$
`yyxxy`	4	$\frac{1}{10}$
`yyxyx`	5	$\frac{1}{10}$
`yyyxx`	6	$\frac{1}{10}$

From this table, we can describe the pmf of W by

w	0	1	2	3	4	5	6
$P(W = w)$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{1}{10}$	$\frac{1}{10}$

R provides a function `dwilcox` to do this pmf calculation for W.

TASK 3. Run the command `dwilcox(0:6, 2, 3)` in R to verify that we got the right probabilities in the W table above.

Now, if $\eta > 0$, then we'd expect more pairs of (X_i, Y_j) with $X_i > Y_j$, and similarly, for $\eta < 0$ we'd expect more pairs with $X_i < Y_j$. So a value of W close to its either extreme (0 or nm) indicates evidence against H_0. So a test for $H_0 : \eta = 0$ can be constructed as:

$$\textit{reject } H_0 \text{ if and only if } W < c_1 \text{ or } W > c_2$$

where $c_1 > 0$ is close to 0 and $c_2 < nm$ is close to nm. This size of this test is $F_W(c_1) + 1 - F_W(c_2)$ where $F_W(w)$ gives the CDF of W when $\eta = 0$ (g does not matter). So to make this test into a size α one, we can choose $c_1 = F_W^{-1}(\alpha/2)$ and $c_2 = F_W^{-1}(1 - \alpha/2)$. These can be calculated in R by using `qwilcox()`.

The following code simulates 10 X_i's from $\mathsf{Normal}(0, 1)$ and 15 Y_j's from $\mathsf{Normal}(0, 1)$ and tests $H_0 : \eta = 0$ with a size 5% test based on W.

```
fn.W <- function(){
        x <- rnorm(10)
        y <- rnorm(15)
        w <- sum(rep(x, 15) < rep(y, each = 10))
        return(w < qwilcox(.05/2, 10, 15) | w > qwilcox(1 - .05/2, 10, 15))
}
```

> TASK 4. What happens when we do `rep(x, k)` for some vector `x` and an integer `k`? What happens if we do `rep(x, each = k)`?

> TASK 5. Use `fn.W` to check if our W-test is size α when g is normal. Then modify and rerun to check it it remains size α if $g = t(1)$.

Power calculation under alternative

Now consider $g = \mathsf{Normal}(0, 1)$. In this case both our t-test and W-test are size 5%. How do they compare in terms of their type II error probabilities? To check this we consider $\eta = 1$ as the alternative point and simulate `x <- rnorm(10) + eta`, `y <- rnorm(15)`.

> TASK 6. Calculate and compare the type II error probabilities of t-test and W-test at $\eta = 1$.