The University of Sydney
School of Mathematics and Statistics

Computer Exercise Week 8

STAT3023: Statistical Inference — Semester 2, 2021

Lecturers: Neville Weber and Michael Stewart

Prepare your computer report and submit to the appropriate Canvas portal by **11:59pm Sunday 10th October**. Please include in your report all the code, plots and any comments required by the questions. The permitted format of the file that you upload will be restricted to PDF, HTML or Word, and should be created in a reproducible fashion (e.g. compiled from an Rmarkdown file). **Please do not have your name visible in your report.**

Two-sided tests for a normal variance

Suppose $X_1, \ldots, X_n$ are iid $N(\mu, \sigma^2)$. In the week 7 Tutorial it was noted that the statistic $Y = (n-1)S^2 = \sum_{i=1}^n (X_i - \bar{X})^2$ (where $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ and S^2 are the sample mean and variance) has a $\sigma^2 \chi^2_{n-1}$ distribution (note we are **not** multiplying by $\frac{1}{2}$ as we did in the week 7 Tutorial!). Consider testing

$$H_0 \colon \sigma^2 = 1 \text{ against } H_1 \colon \sigma^2 \neq 1. \tag{1}$$

1. One possible level-α test is the "equal-tailed" test based on Y, where we reject for $Y < a$ or $Y > b$ where

$$P_0\{Y < a\} = P_0\{Y > b\} = \frac{\alpha}{2}.$$

(a) Taking $\alpha = 0.04$ and $n = 5$, find appropriate values `a` and `b`.

(b) Defining `sig.sq=(50:150)/100` plot the power of the test against `sig.sq`. Add a horizontal dotted line to indicate the level.

2. In Tutorial week 7 we also saw that the UMPU test rejects for large values of

$$S^2 - \log(S^2)$$

which is equivalent to rejecting for small values of the statistic

$$T = (n-1)\log Y - Y\,;$$

to see this, write $\log(S^2) = \log Y - \log(n-1)$, multiply through by $n-1$ and ignore the $(n-1)\log(n-1)$ term.

If the test is to have level α, we reject for $Y \leq c$ or $Y \geq d$ where

$$P_0(Y \leq c) + P_0(Y \geq d) = \alpha \tag{2}$$

and

$$(n-1)\log(c) - c = (n-1)\log(d) - d\,. \tag{3}$$

(a) Write a function of the form

```r
fn=function(c,alpha,n) {
    ...
}
```

which

- computes the appropriate `d` so that `c` and `d` satisfy (2);
- then computes and outputs the difference between the left-hand side and right-hand side in (3).

(b) Use the R function `uniroot()` to find the root (in `c`) of the equation `fn(c,0.04,5)=0`. In your code you will need a command along the lines of

```
uniroot(fn,lower=0,upper=...,alpha=0.04,n=5)
```

Consult the week 7 exercise for some hints as to how to choose the `upper=...`. When you have worked out the right commands, wrap it all in a function of the form

```
norm.var.umpu=function(alpha,n) {
    ...
}
```

which returns a list containing elements `$c` and `$d`.

(c) Recreate your plot from part (b) of the previous question and add to it the power function of the UMPU test.

3. The GLRT test of (1) above uses the statistic

$$L_n = \ell(\bar{X}, Y/n; \mathbf{X}) - \ell(\bar{X}, 1; \mathbf{X}) = -\frac{n}{2}\log\left(\frac{Y}{n}\right) - \frac{n}{2} + \frac{Y}{2}$$

which is an increasing function of $Y - n\log Y$ (as opposed to the UMPU which rejects for large $Y - (n-1)\log Y$). Adapt your code for the previous question to compute the power of the *exact* GLRT, recreate your earlier plot and add a power curve to it so it shows all 3 power curves on the 1 graph. Add an informative heading, legend, etc.. Comment on the main differences between the 3 tests.

4. As a final step, recreate your last plot but use an extended range for the parameter: `sig.sq=(1:400)/100`.