Hypothesis testing

This assessment requires you to use R to perform some hypothesis testing and work with some properties of the Gaussian distribution. Each question is of equal mark value.

First, execute the following code which will give you a mean and standard deviation. Students who use my ID will receive zero credit.

```
set.seed(1266402)      # This is my student ID: replace with yours!
rnorm(1,10,2)          # Use this for the mean
```

```
## [1] 11.40135
```

```
sqrt(rchisq(1,df=2))  # Use this for the standard deviation
```

```
## [1] 0.7041335
```

Gaussian distribution

Above we see that the mean of the distribution is $\mu = 11.40$ and the standard deviation is about 0.704. But use your values, not mine. For the questions below, use the course manual as a guide.

- Plot the probability density function for your distribution
- If random variable X is drawn from your Gaussian distribution, find the probability that X lies in the range $(\mu - \sigma, \mu + 2\sigma)$
- Use `rnorm()` to generate a random sample from your distribution and give a histogram of your observations
- For bonus credit, find the steepest point of the standard normal distribution (this is a numerical exercise, not calculus).

Student t-test

The average weight of a banana is 100 grams. An agricultural scientist buys bananas from a supermarket. Their weight, in grams, is as follows:

```
w <- c(102.5, 102.9, 99.9, 101.9, 101.3, 98.9, 96.5, 97.7, 101.8,
100.3)
```

She suspects that this sample of bananas is heavier than average and wonders if this supermarket is selling bananas that are heavier than the notional 100g mean weight.

- State a sensible null hypothesis
- State the precise definition of p-value and explain what "more extreme'' means in this context
- Is a one-sided or two-sided test needed? justify
- Perform a student t-test using R and interpret
- Give a 95% and 99% confidence interval for the mean.
- For bonus credit, perform a Z test and account for any differences you find

Hypothesis testing: binomial distribution

A NZ scientist is studying a particular type of plant that always has exactly 7 flowers on it. Each flower may be fertilized or unfertilized. The probability of a flower being fertilized is known to be 0.4; flowers are independent of one another and you may use the binomial distribution.

- Using `dbinom()` or otherwise, what is the probability that exactly one flower on a particular plant is fertilized?
- What is the probability that three or more flowers are fertilized?
- A plant is found to be very near a beehive, and it is suspected that bees fertilize this kind of plant. The scientist finds that 5 flowers are fertilized.
- State a sensible null hypothesis
- state the precise definition of p-value and define what "more extreme" means in this context.

- One-sided or two sided? Justify
- Calculate the p-value. Is it significant?
- Interpret your finding in a way that a busy entomologist, who is not a statistician, could understand