MAT 5030: Homework 3

- Assigned on October 17 (Mon), 2016.
- **Due at 5:30pm on October 24 (Mon)**, 2016.
- Submit **by e-mail at seikibunpu@gmail.com**
- The e-mail title will be "M5030 HW3 - Your Name".
- Use the following format.
  - R-code, output, and brief explanations in one TXT or MS-Word file. Cut and paste as text. Do not use screenshots.
  - Graphics in **ONE** MS-Word or PDF file.
  - In case of MS-Word, you can put both in one file.
- Truncate the output if a single output exceeds 20 lines in the R console.
- A late submission will reduce your score as follows. E.g, if you submit 19 hours late and your work gets 9 points out of 10, your score will be $9 \times 0.8 = 7.2$ out of 10.

Delayed by (in days)	0-0.5	0.5-1	1-1.5	$\cdots$	4-4.5	4.5-
Multiplier	0.9	0.8	0.7	$\cdots$	0.1	0

1.

(a) Make a function to simulate $E[Z^k]$ where Z is a standard normal random variable and $k = 0, 1, 2, 3, \cdots$. The inputs are the exponent k and the sample size n of simulation. Namely given k and n, generate standard normal random numbers $z_1, \cdots, z_n$ and estimate $E[Z^k]$ by

$$\frac{1}{n}\sum_{i=1}^{n} z_i^k.$$

Set $n = 100,000$ as the default value. For example, we expect the following output (the result varies slightly due to different random numbers).

```
> normmoment(4) # E[X^4] based on 100,000 obs.
[1] 2.971398
> normmoment(4,n=500000) # E[X^4] based on 500,000 obs.
[1] 3.008832
```

(b) Apply the function in (a) for $k = 2, 4, 6, 8$ and $n = 500,000$. (Note: The answers are not unique, and depend on simulation.)

2. (Wilcoxon signed-rank test, Exercise 1 in Chapter 5 Slides p.36)

Do a simulation to calculate $P(V \leq 9)$ when the sample size is 7. To be concrete, repeat a procedure to generate 7 random numbers from a symmetric distribution with mean zero and calculate V for 10,000 times. Report the sample proportion that $V \leq 9$. (You can use the '`wilcox.test`' function as necessary).

3. (Test for equal variance; Welch's two-sample t-test)

The exam scores of 19 students (12 male and 7 female) are as follows. Assume that exam scores are normally distributed for each gender group.

```
> Male
 [1] 41 51 59 38 58 55 56 72 37 74 44 38
> Female
[1] 54 62 61 58 52 55 70
```

(a) Make a boxplot by gender in one pane.

(b) Test if the sample variance for male and the sample variance for females are different (You may use any R functions in the default package). Report the sample variances, F-statistic and p-value. Is the null hypothesis rejected at 5% significance level? How about at 10% significance level?

(c) Implement a Welch's two-sample t-test to see if two genders have different sample means by using the `'t.test'` function.

(d) Implement the same t-test as (b) without using the `'t.test'` function. You can use basic operations and functions (`+, -, *, /, ^, mean, sd, pt, qt, sqrt, min, max`). Report the t-statistic, the number of degrees of freedom and p-value.

Note: All tests in Question 3 are two-sided.