Homework 1

Instructions:

- Homeworks will not be collected nor graded, but it will be the basis for quizzes and exams.

Reading: Ch1 and Appendix A and B of textbook.

1. A hospital administrator codes incoming patients suffering gunshot wounds according to whether they have insurance (coding 1 if they do and 0 if they do not) and according to their condition, which is rated as good (g), fair (f), or serious (s). Consider an experiment that consists of the coding of such a patient.
   (a) Give the sample space of this experiment.
   (b) Let A be the event that the patient is in serious condition. Specify the outcomes in A.
   (c) Let B be the event that the patient is uninsured. Specify the outcomes in B.
   (d) Give all the outcomes in the event $B^c \cup A$.

2. A salesperson has scheduled two appointments to sell encyclopedias. Her first appointment will lead to a sale with probability .3 and her second will lead independently to a sale with probability .6. Any sale made is equally likely to be either for the deluxe model, which costs \$1000, or the standard model, which costs \$500. Determine the probability mass function of X, the total dollar value of all sales. That is, what are the possible values of X, and their associated probabilities?

3. Suppose that the distribution function of X is given by

$$F(b) = \begin{cases} 0, & b < 0, \\ \frac{b}{4}, & 0 \le b < 1 \\ \frac{1}{2} + \frac{b-1}{4}, & 1 \le b < 2 \\ \frac{11}{12}, & 2 \le b < 3 \\ 1, & b \le 3. \end{cases}$$

   (a) Find $P(X = i), i = 1, 2, 3$.
   (b) Find $P(1/2 < X < 3/2)$.

4. Suppose X is a random variable with probability density function

$$f(x) = \begin{cases} C(4x - 2x^2), & 0 < x < 2, \\ 0, & \text{otherwise}. \end{cases}$$

(a) Find C to be a density function.

(b) Compute $E(X^{-1})$.

5. The joint probability density function of X and Y is given by

$$f(x, y) = \frac{6}{7}\left(x^2 + \frac{xy}{2}\right), \quad 0 < x < 1, 0 < y < 2$$

   (a) Verify that this is indeed a joint density function.
   (b) Compute the density function of X.
   (c) Find $P(X > Y)$.
   (d) Find $P(Y > 1/2 | X < 1/2)$.
   (e) Find $E(Y)$

6. Show that if $\mathbf{X}$ is an n-dimensional random vector such that $\mathbf{X} \sim MVN(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, and $\mathbf{B}$ is a real $m \times n$ matrix, and $\mathbf{a}$ is a real m-vector, then

$$\mathbf{Y} = \mathbf{a} + \mathbf{BX} \sim \mathbf{MVN}(\mathbf{a} + \mathbf{B}\boldsymbol{\mu}, \mathbf{B\Sigma B'})$$

7. In this problem we will see that long-run relative frequency converges to $P(A)$. Suppose that you toss a coin and interested in the event $A = \{H\}$ with $P(A) = 1/3$. The following R code generates 10000 independent Bernoulli random variables with success probabilities $p = 1/3$. Then, you calculate relative frequency and plot them. Does relative frequency converge to some constant (existence)? Repeat the same experiment again and confirms that they always converges to the same number (uniqueness).

```
n=10000;
pr = 1/3
y = rbinom(n, 1, prob=pr);
rel.freq = cumsum(y)/(1:n);
plot(rel.freq, type="l", ylim=c(0,1));
abline(a=pr, b=0, col="red")
title("Long run relative frequency");
```

   R can be downloaded from http://www.r-project.org/. If you are new to R, see also "An introduction to R" under manuals tab.

8. (Special case of Problem 1.10 in the textbook, p. 42.) If $m_t = a + bt + ct^2$ is a polynomial of degree 2, show that ∇m_t is a polynomial of degree 1, $\nabla^2 m_t$ is a polynomial of degree 0 and $\nabla^3 m_t = 0$.

9. Problem 1.11 Do only part (a)

10. Problem 1.12

11. (Cauchy-Schwarz inequality) Prove Cauchy-Schwarz inequality given in the following various forms.

   (a) Suppose that $x_1, \dots, x_n \in \mathbb{R}$ and $y_1, \dots, y_n \in \mathbb{R}$. Then, show that

   $$|x_1 y_1 + \ldots + x_n y_n|^2 \leq (x_1^2 + \ldots x_n^2)(y_1^2 + \ldots + y_n^2). \tag{1}$$

   This is what you have learned in high school.

   (b) Find the definition of inner-product space over the real field (real number $\mathbb{R}$). (Find any linear algebra book)

   (c) In fact, CS inequality in (1) can be generalized to **any** inner-product. That is, CS inequality in (a) can be rewritten as the following. Consider the inner-product in the Euclidean space $\mathbb{R}^n$ defined as

   $$\langle x, y \rangle = x' y = x_1 y_1 + \ldots + x_n y_n,$$

   where $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ and $y = (y_1, \dots, y_n) \in \mathbb{R}^n$. Then, (1) is equivalent to

   $$|\langle x, y \rangle|^2 \leq \langle x, x \rangle \langle y, y \rangle.$$

   In statistics, we define inner-prduct as $\langle Z, W \rangle = E(ZW)$. Hence, the CS inequality becomes

   $$|\mathrm{Cov}(Z, W)| \leq \sqrt{\mathrm{Var}(Z)\,\mathrm{Var}(W)} \tag{2}$$

   by taking $Z = X - E(X)$ and $W = Y - E(Y)$. Show that (2) holds.

   Hint. For any inner product, use that

   $$\langle x - ty, x - ty \rangle \geq 0 \quad \text{for all } t.$$

12. (Non-negative definiteness) In matrix algebra, $n \times n$ matrix Γ is called non-negative definite if

   $$a' \Gamma a \geq 0$$

   for any vectors $a = (a_1, \dots, a_n)'$ with real entries. Determine whether the following matrices are non-negative definite or not.

   $$\Gamma_1 = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad \Gamma_2 = \begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix}.$$

   Remark. In our class, we have seen that $\gamma(\cdot)$ is a autocovariance function of a **stationary time series** if and only if it is *even* and *non-negative definite*. In light of above definition of non-negative definiteness, we defer that Γ is a autocovariance of a **stationary time series** if it is symmetric (because it is even), that is, it is written as

   $$\Gamma = \begin{pmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(n-1) \\ \gamma(1) & \gamma(0) & \cdots & \gamma(n-2) \\ \vdots & \vdots & \vdots & \vdots \\ \gamma(n-1) & \cdots & \cdots & \gamma(0) \end{pmatrix}$$

and non-negative definite, that is,

$$a'\Gamma a \geq 0$$

for any vector $a = (a_1, \dots, a_n)'$.

13. Problem 1.4

14. Problem 1.8 Do only part (a). **Hint.** To show that $\{X_t\}$ is WN(0,1) you must show that *i*) $EX_t = 0$ for all t and *ii*) $\text{Var}X_t = 1$ for all t and *iii*) $\text{Cov}(X_{t+h}, X_t) = 0$ if $h \neq 0$ for all t.

15. In this problem you are asked to do preliminary real data analysis. Open data set IMPORTS.TXT attached.

    (a) Produce a time plot of the time series and discuss its features (such as trend, seasonality, outliers, etc.).

    (b) (Regression) Fit polynomial regression to remove trend. You need to determine the order of polynomial fit by looking at fitted regression line and examining residuals. State the final regression model you have selected.

    (c) (Smoothing) Apply 5 point moving average to remove trend. Similar to (b), discuss result by overlaying fitted trend and examining residuals. Use `smooth.ma` function provided in itsmr package.

    (d) (Smoothing) Apply exponential smoothing with $a = .4$ and discuss result as in (b). Use `smooth.exp` function provided in itsmr package.

    (e) (Differencing) Remove trend by applying differencing. Which order gives you the best fit? Explain your reasoning. To apply (lag 1) differencing in R, do `diff(data)`.

    (f) Draw correlogram from the data set. Briefly describe the features of correlogram.

    (g) Draw correlogram from the residuals obtained after a linear trend fit. (This is done in part (b)). Describe its feature. Provide results of the tests of randomness with a discussion of Ljung - Box, McLeod - Li, Turning points tests (that is, whether IID noise is rejected according to these tests); is IID model rejected based on the correlogram? In R, use `test()` function provided in `itsmr` package.

    (h) Draw correlogram from the residuals obtained after applying 5 point moving average. (This is done in part (c)). Describe its feature. Provide results of the tests of randomness with a discussion of Ljung - Box, McLeod - Li, Turning points tests (that is, whether IID noise is rejected according to these tests); is IID model rejected based on the correlogram?