CAS CS 365
Lab 4

1. Covariance of X, Y
    a. A measure of how well X and Y vary together
    b. We are often interested in two or more random variables at the same time
    c. Consider the following examples:
        i. The relationship between height and weight
        ii. The frequency of exercise and the rate of heart disease
        iii. Air pollution levels and rate of respiratory illness
2. Covariance computation from joint pdf
    a. Let's say we are given the following information
    b. Le X and Y be continuous random variables with joint pdf

$$f_{X,Y}(x,y) = 3x, \qquad 0 \leq y \leq x \leq 1,$$

and zero otherwise.

i. Compute Covariance of X, Y

$$Cov_{f_{X,Y}}[X,Y] = E_{f_{X,Y}}[XY] - E_{f_X}[X] E_{f_Y}[Y]$$

Expected Value of XY with joint probability density function (pdf), f of X,Y

Expected value of X with pdf, f of X

1. Expected value of Y with pdf, f of Y

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = \int_0^x 3x dy = 3x^2, \qquad 0 \leq x \leq 1,$$

$$E_{f_X}[X] = \int_{-\infty}^{\infty} x f_X(x)dx = \int_0^1 x \times 3x^2 dx = \left[\frac{3}{4}x^4\right]_0^1 = \frac{3}{4},$$

2.

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx = \int_y^1 3x dx = \left[\frac{3}{2}x^2\right]_y^1 = \frac{3}{2}(1-y^2), \quad 0 \leq y \leq 1,$$

$$E_{f_Y}[Y] = \int_{-\infty}^{\infty} y f_Y(y)dy = \int_0^1 y \times \frac{3}{2}(1-y^2)dy = \left[\frac{3}{2}\left(\frac{y^2}{2} - \frac{y^4}{4}\right)\right]_0^1 = \frac{3}{2}\left(\frac{1}{2} - \frac{1}{4}\right) = \frac{3}{8},$$

3.

4.

$$
\begin{aligned}
E_{f_{X,Y}}[XY] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{X,Y}(x,y)dydx = \int_0^1\int_0^x xy \times 3xdydx \\
&= \int_0^1 \left\{\int_0^x ydy\right\} 3x^2dx = \int_0^1 \left[\frac{y^2}{2}\right]_0^x 3xdx = \int_0^1 \frac{x^2}{2} \times 3x^2dx \\
&= \frac{3}{2}\left[\frac{x^5}{5}\right]_0^1 = \frac{3}{10},
\end{aligned}
$$

5.

$$
Cov_{f_{X,Y}}[X,Y] = E_{f_{X,Y}}[XY] - E_{f_X}[X]E_{f_Y}[Y] = \frac{3}{10} - \frac{3}{4} \times \frac{3}{8} = \frac{3}{160}
$$

3. Remarks
   a. Events and random variables
   b. PDF of continuous random variables
   c. Chebyshev's inequality

Let $X_1, X_2, \ldots$ be a sequence of iid RVs with mean μ, and standard deviation σ.
Consider the sum $S_n = X_1 + \ldots + X_n$

$$
\Pr\left(\left|\frac{S_n}{n} - \mu\right| > \epsilon\right) \le \frac{\mathbb{V}\mathrm{ar}(S_n/n)}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2}
$$