Session 7: Probability Review II

You can use Python to help crunch numbers in this Handout.

Q1 (Investment Portfolio)

An investment consultant believes that the probability distribution of the return on investment (in percent per year) on a certain portfolio is given in the below table.

Returns on investment	Probability
9	0.07
10	0.15
11	0.23
12	0.25
13	0.15
14	0.12
15	0.03

a) What is the probability that the porfolio's return will be at least 12%?

b) What is the probability that the return will be between 10% and 13% (inclusive)?

c) What is the expected return?

d) What is the variance and standard deviation of the returns?

Q2 (Quality Control)

A factory manufactures 200 units per hour. According to past data, each unit is of good quality with probability 0.95, independent of other units.
a) What is the expected number of bad quality units produced in a 4 hour shift?

b) What is the standard deviation of the number of bad quality units produced in a shift?

c) What is the probability that there are (strictly) less than 30 bad units produced in a shift?

d) What is the probability that there are between 30 and 50 bad units produced in a shift (inclusive)?

e) What is the probabiltiy that (strictly) more than 55 bad units produced in a shift?

Q3 (Sports Analytics)

The men's 100 meter sprint at the 1996 Olympic Games in Atlanta was a hotly contested event between Donovan Bailey of Canada, Frankie Fredericks of Namibia, and Ato Boldon of Trinidad. Assume that for all three runners, the probability distribution of the time to run the race a (continuous) uniform distribution between 9.75 and 9.95 seconds, and that the running times are independent.

a) What is the probability that Donovan Bailey's time will beat the previous record of 9.86 seconds?

b) What is the probability that the winning time will beat the previous record of 9.86 seconds?

Central Limit Theorem (Normal Approximation): If $X_1, X_2, \cdots X_n$ are independent random variables and n is "large," then the sum $Y = X_1 + X_2 + \cdots + X_n$ is approximately normal distributed, with

$$\mu = E[X_1] + E[X_2] + \cdots + E[X_n]$$
$$\sigma^2 = Var(X_1) + Var(X_2) + \cdots + Var(X_n)$$

Q4 (Demand Estimation)

A marketing team would like to estimate the demand for a new product before its launch. They have segmented the market into 4 segments. The number of customers of each segment and the estimated purchase probabilties are as below. (Assume that each potential customer makes their purchase decision independently.)

Segment	# of Customers	Purchase Probability
A	10000	0.1
B	5000	0.3
C	300	0.9

a) What is the mean and standard deviation of the total demand for the product?

b) Use the normal approximation to estimate the probability that demand is less than 2500.

c) Use the normal approximation to estimate the probability that the demand is more than 2800.