INF389 Statistical methods in bioinformatics (H2003)
Exercise 1

Eivind Coward

For 20 Oct

Problem 1 (1.24 from Ewens & Grant)

Two fair coins are tossed and you are told that at least one coin came up heads. What is the probability that both coins came up heads? Note: the two "instinctive" answers $\frac{1}{2}$ and $\frac{1}{4}$ are both incorrect. It is necessary to use the formula for conditional probability.

Problem 2 (Discrete distribtions)

A *Bernoulli trial* is a single trial with two outcomes, called "success" and "failure". The probability of success is denoted p.

a) Let Y be a random variable corresponding to the number of successes after 1 Bernoulli trial. What is the (point) probability distribution $P_Y(y)$?

b) Calculate the mean and variance of Y.

c) A binomial random variable Z with parameters n and p can be defined as the number of successes in n independent Bernoulli trials with success probability p. Show that this leads to the probability distribution given by

$$P_Z(z) = \binom{n}{z} p^z (1-p)^{n-z}, \quad z = 0, 1, \ldots, n$$

d) Find the mean and variance of Z.
Hint: Use the definition, not the distribution formula!

Problem 3 (Continuous distributions)

Let X be a random variable with the uniform distribution, given by the probability density function (pdf)

$$f_X(x) = \frac{1}{b}, \quad 0 \leq x \leq b \tag{1}$$

a) Calculate the distribution function $F_X(x) = \Pr(X \leq x)$.

b) Find the mean and variance of X.

c) Now let X and Y be two independent identically distributed (iid) random variables, also having the uniform density function (1). Write down their joint density function $f_{X,Y}(x, y)$.

d) In general, the pdf of the sum $Z = X + Y$ of two independent random variables X and Y is given by their *convolution*

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(z - y) f_Y(y)\, dy$$

Can you explain why this makes sense? Calculate $f_Z(z)$ when X and Y are as above.

e) What is the probability that Z is between $b/2$ and $3b/2$?

f) Find the mean and variance of Z.