Biostatistics
Practice Exercises on Probability

Instructor: Sophia Daskalaki

Indicative Reading Assignment:

1. Biostatistics: A Foundation for Analysis in the Health Sciences, by W.W. Daniel and C.L.Cross

   Ch. 1, 3 and 4

2. Probability & Statistics for Engineers & Scientists, by Walpole, Myers, Myers, Ye

   Ch.2 (except 2.3), Ch.3, Ch.4 (except 4.4), 5 (except multinomial distibution, 5.3, 5.4), 6.2, 6.3, 6.4, and exponential from 6.6

3. From Engineering Biostatistics, Ch. 3 (except 3.5 and 3.9), 4 (except 4.4), 5 and 6 (from the distributions only those covered in class)

- *From the second book you should be able to work out the exercises*: 2.76, 2.78, 2.85, 2.93, 2.94, 2.118, 2.120, 3.3, 3.8, 3.23, 3.30, 4.4, 4.13, 4.22, 4.39.
- *From the third book you should be able to work out the exercises*: 4.3, 4.4. You may also try several others in Ch. 4
- as well as the exercises listed below:

Exercise 1. There is a $50-50$ chance that a queen carries the gene of hemophilia. If she is a carrier, then each prince has a $50-50$ chance of having hemophilia independently. If the queen is not a carrier, the prince will not have the disease. Suppose the queen has had three princes without the disease. What is the probability the queen is a carrier? Comment on the difference you observe between the *a priori* probability that the queen is a carrier and the corresponding *a posteriori* probabilities knowing that the princes are disease free.

Exercise 2. According to birth statistics dizygotic twins have the same probability of each gender as in the rest of all births, which assume is approximately 52% male and 48% female. On the other hand monozygotic twins are always of the same gender. It is estimated that among all twin pregnancies, about 1/8 are monozygotic. In a twin pregnancy find the probability of two boys if it is

(a) monozygotic pregnancy,

(b) dizygotic pregnancy,

(c) dizygotic pregnancy given that we know that the gender of the babies is the same.

If Mary is expecting twins, but no information about the type of pregnancy is available, what is the probability that the babies are:

(d) two boys,

(e) of the same gender;

(f) Find the probability that Mary's pregnancy is dizygotic if it is known that the babies are two girls.

Exercise 3. Let E and F be mutually exclusive events in the sample space of an experiment. Suppose that the experiment is repeated until either event E or F occurs. Show that the probability that event E occurs before event F is $P(E)/[P(E) + P(F)]$

Exercise 4. A "random walker" takes a step (of a fixed length L) to the left or to the right of his/her position with equal probabilities.

(a) Write down the probability distribution of the person's position starting at $m = 0$ after one single step, two steps, three steps, or four steps.

(b) What is the mean of each of the distributions in question (a)?

(c) What is the variance of each of the distributions in question (a)?

Exercise 5. Consider the following life table that refers to a certain population.

Number of men remaining alive at ten-year intervals			
Age in years, x	Number Surviving, l_x	Age in years, x	Number Surviving, l_x
0	1000	60	758
10	959	70	524
20	952	80	211
30	938	90	22
40	920	100	0
50	876		

This table shows the number of men, from a group numbering 1000 at birth, who we would expect to be alive at different ages. Thus, for example, we expect that after 10 years, 959 survive and 41 have died, at 20 years 952 survive and so 7 more have died.

(a) Calculate the probabilities that a randomly chosen individual will survive to ages 10, 20, 30, 40, 50, 60, 70, 80, 90, 100. Does this set of probabilities form a probability mass function? Justify your answer.

(b) What is the probability that a randomly chosen individual aged 60 years will survive to age 70?

(c) What is the probability that two randomly chosen individuals aged 60 years both will survive to age 70?

(d) For each decade, what is the probability that a randomly chosen individual dies in that decade? Does this new set of probabilities form a probability mass function? Justify your answer.

(e) As an approximation, we can assume that the average number of years lived in the decade of death is 5. Thus, those who die in the 2nd decade will have an average life span of 15 years. According to such an assumption and the results of this study what is the mean lifetime of all men? (This is the expectation of life at birth.)

Exercise 6. According to a study the anti-inflammatory drug Piroxicam may cause serious gastrointestinal reactions to some individuals. In fact, it is known that 538 serious gastrointestinal reactions were reported to the British Committee on Safety of Medicines out of 9,160,000 prescriptions. If we can assume that serious gastrointestinal reactions to Piroxicam follow a Poisson process, calculate:

(a) the rate of gastrointestinal reactions per $10,000$ prescriptions

(b) the probability of exactly two gastrointestinal reactions per $10,000$ prescriptions

(c) the probability of finding at least two gastrointestinal reactions per $10,000$ prescriptions.

Exercise 7. Given the normally distributed variable X with mean 18 and standard deviation 2.5, find the probabilities:

(a) $P(X < 15)$

(b) $P(17 < X < 21)$

(c) the value of k such that $P(X < k) = 0.2236$;

(d) the value of k such that $P(X > k) = 0.1814$;

Exercise 8. Leptin is a hormone that plays a key role in regulating energy intake and energy expenditure, including the regulation of appetite and metabolism. Serum leptin concentrations can be measured in several ways and one approach is by using a radioimmunoassay in venous blood samples. Several studies have consistently found women to have higher serum leptin concentrations than men. For the adults of a certain country, across a broad age range, the serum leptin concentration in women is approximately normal with mean 12.7 $\mu g/L$ and standard deviation 1.14 $\mu g/L$ and in men approximately normal with mean 4.6 $\mu g/L$ and standard deviation 0.7 $\mu g/L$.

(a) What is the probability that the concentration of leptin in a randomly selected adult male in this country exceeds 6 $\mu g/L$?

(b) What proportion of women have concentration of leptin in the interval $12.7 \pm 2\ \mu g/L$?

(c) What interval, symmetric about the mean $12.7\ \mu g/L$, contains the leptin concentrations of 95% of all adult women?

(d) What is the probability that two randomly selected adults, one male and one female of this country, will differ by less than $2.5\ \mu g/L$ in their concentration of leptin measured in their blood.