BIOE: Biostatistics Course Fall 2017

Assignment 2

Due 19th October

For plotting use Matplotlib and provide the code you used with the assignment.

1. Which of the following adequately describes a random variable:

a) All possible outcomes of an experiment
b) The measurement of one experimental outcome
c) The probability of an experimental outcome
d) The frequency of an outcome

2. You have engineered a cardiac muscle cell with a florescent marker. You measure the fluorescence in 20 different cell cultured under identical conditions using relative fluorescence units. If x is the degree of fluorescence, what type of variable is x?

a) x is a discrete random variable
b) x is a continuous random variable
c) x is a constant
d) x is not a random variable

3. State what the symbol $p_X(x)$ means

4. Sketch the probability mass function for the following data:

x	1.1	2.5	4.1	4.6	5.3	6.1
$p(x)$	0.16	0.14	0.11	0.27	0.22	0.1

5. Provide reasons why the following discrete probability distributions are not valid PMFs:

x	0.5	0.25	0.25
$p(x)$	-0.4	0.6	0.8

x	1	2	4	5	6
$p(x)$	0.16	0.14	0.11	0.27	0.22

6. Given the following probability distribution for the random variable X:

x	13	18	20	24	27
$p(x)$	0.22	0.25	0.2	0.17	0.16

compute the following:

a) $p(18)$
b) $P(X > 18)$
c) $P(X \leq 18)$
d) Mean, μ of X
e) Variance, σ^2 of X
f) Standard deviation, σ of X

7. In class we briefly introduced the cumulative distribution function for continuous random variables. A similar idea exists for discrete random variables. Research this topic in your textbooks or the internet and fill in the missing entries from the cdf row for the following discrete distribution. Note the symbol for the cdf is $F(x)$

Missing entries are indicated with a question mark. Although the distribution clearly goes on, you only need to fill in the entries up to 8.

X	1	2	3	4	5	6	7	8	...
$p(x)$	0.3	0.21	0.147	0.103	0.072	0.0504	0.0354	0.025	...
$F(x)$	?	?	?	?	?	?	?	?	

8. Plot $p(x)$ and $F(x)$ from the previous question as a function of the random variable X. Plot both on the same graph. Only plot up to $X = 8$.

9. The geometric distribution has the form $(1-q)^{k-1}q$. Show that the mean for a geometric distribution is $1/q$. You may research your answer from textbooks or the internet.

10. A pharmaceutical company is looking for a new drug to combat sleeping sickness. They have 20 drug candidates to test. The probability that a drug will

be effective is 0.05. Let X denote the random variable that describes the number of drugs tested until they find a success.

a) What is the probability that the company must select 4 drugs before they find a drug that is effective?

b) What is the average number of drugs the company much check before they find an effective drug?

11. Let X be a continuous random variable whose probability density function is

$$f(x) = \begin{cases} 2x & 0 \leq x \leq 1 \\ 0 & otherwise \end{cases}$$

a) Sketch the probability density function $f(x)$.

b) Calculate the probability that X falls between 0 sand 1/2

c) Calculate the probability that X falls between 1/4 sand 3/4

12. On average two people suffer from extremely bad strokes per year in Washington state. What is the probability that five people will suffer bad strokes next year?

You can assume that the distribution of strokes follows a Poisson distribution. A Poisson distribution is given by:

$$p(k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

where λ equals the average number of events (strokes) per interval (in this case per year), and k is the number of events in an interval, ie $p(k)$ is the probability of k events (strokes) happening in a year. Note that the Poisson distribution is a discrete distribution.

a) Plot the Poisson distribution, assuming $\lambda = 2$. Make a reasonable judgment on the length of the x axis to make the graph look sensible, ie size of k.

b) What is the probability that five people will suffer bad strokes next year in the state?

13. **UNDERGRADUATES ONLY**

Given a random variable, X we know that multiplying the random variable by a constant causes the variable to go up by the square of the constant, ie:

$$Var(aX) = a^2 Var(X)$$

Write a computer program to show numerically that this is indeed the case. Show numerically that this relationship is independent of the sample size of X.

14. **GRADUATES ONLY**

Let X be a continuous random variable whose probability density function is

$$f(x) = \begin{cases} 2(1-x) & 0 \leq x \leq 1 \\ 0 & otherwise \end{cases}$$

Show that this distribution is a legitimate PDF by:

a) showing that $f(x)$ is always positive

b) showing that the area under $f(x)$ is one.

c) Calculate the mean of X

d) Calculate the standard deviation of X