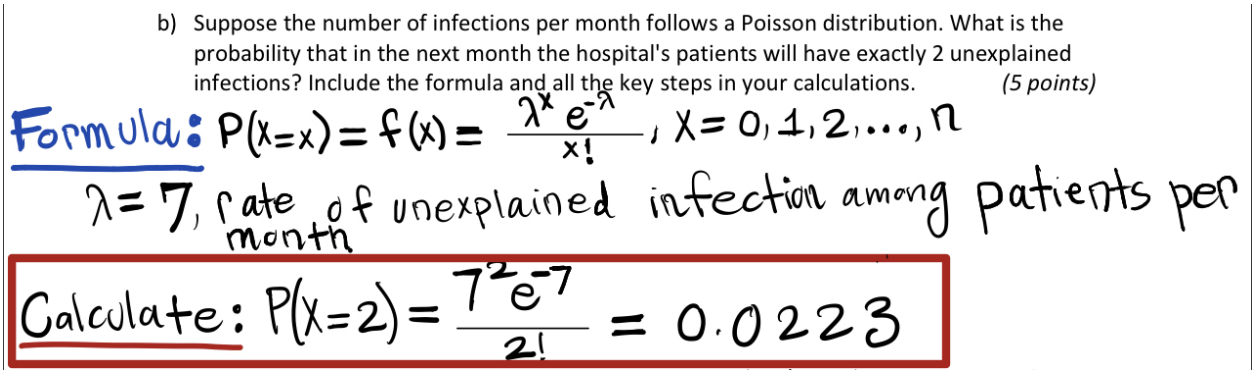
BiostatisticsMidterm-tk2886

Problem 1

a). Assumptions that must be true to use the Poisson Distribution to model the number of infections per month:

1. Events occur one at a time; two or more events cannot occur exactly at the same time and location;
2. The occurrence of an event in a given period is independent of the occurrence of an event in a non-overlapping period;
3. The expected number of events during any period is constant.

b).



We may use R code to get the same value.

```
prob = dpois(2, 7)
prob
```

```
## [1] 0.02234111
```

The probability that in the next month the hospital's patients will have exactly 2 unexplained infections is: **0.0223411**.

Problem 2

Formula: Bayes Theorem

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{P(A|B_i)P(B_1) + P(A|B_2)P(B_2) + \dots P(A|B_k)P(B_k)}$$

D = developed CHD

C = Initial serum cholesterol levels above 200

$P(D) = 0.25$, $P(D^c) = 1 - 0.25 = 0.75$

$P(C|D) = 0.60$

$P(C|D^c) = 0.16$

Interested in:

$$P(D^c|C^c) = \frac{P(D^c \cap C^c)}{P(C^c)} = \frac{P(C^c|D^c)P(D^c)}{P(C^c|D^c)P(D^c) + P(C^c|D)P(D)}$$

$P(C^c|D^c) = 1 - P(C|D^c) = 0.84$

$P(C^c|D) = 1 - P(C|D) = 0.40$

$$\frac{(.84)(.75)}{(.84)(.75) + (.40)(.25)} = \frac{.63}{.63 + .1} = 0.8630 \approx 0.86$$

The probability that a random chosen subject will not develop CHD, given that he had an initial serum cholesterol level below or equal 200 is: **0.86**.

Problem 4

```
dbinom(4, 10, 0.5) + dbinom(6, 10, 0,5)
```

```
## [1] -Inf
```