Midterm Exam

Name:________________________

1. **(15 points)** You are picking up snacks for 12 friends. You want each person to have their own snack. That is to say that none of your friends share snacks. You get to the store and there are only 3 kinds of snacks available. How many ways can you assign snacks to your friends if

   (a) The store has an unlimited stock of each kind of snack

   3^12 ways=531441

   (b) The store only has 4 packages of each kind of snack

   There are essentially 12 items to distribute among 12 people but there are 4 duplicates of each of the three kinds. So, the number of ways to do this is
   12!/(4!*4!*4!) = 34650

2. **(10 points)** Suppose there are only two illnesses going around with 80% of people who are sick having the common cold and 20% having the flu. Additionally, assume that nobody has both illnesses at the same time. A symptom of the flu is a fever and the probability of a person having a fever if have the flu is 0.9. On the other hand, the probability a person having a fever if they have the cold is 0.15.
   You wake up one day with a fever. What is the probability that you have the cold?

   Let us denote the cold/flu as C/V and the event of having a fever as F.

   We have from the question that P(F|V)=0.9, P(F|C)=0.15, P(C)=0.8 and P(V) = 0.2

   P(C|F) = P(F|C)P(C)/P(F)
   = P(F|C)P(C)/(P(F|C)P(C) + P(F|V)P(V))
   = 0.15*0.8/(0.15*0.8+0.9*0.2)
   = 0.4.

3. **(15 points)** Sequencing technologies can "read" many short fragments (simply called reads) from a genome. Given that the process through which the read sequences are generated is random, it is possible that certain parts of the genome will remain uncovered unless an impractical amount of sequences are generated. Human genome is 3×10^9 bp long. A patient's genome has been sequenced and it is randomly covered by 300×10^5 reads (each read is 100 bp long). We assume that the number of times a base in the human genome is covered follows a Poisson distribution.

   (a) **(5 points)** What is the probability that a particular base is not covered by any reads?

   Answer: $\lambda = \dfrac{300 \times 10^5 \times 100}{3 \times 10^9} = 1 \qquad P(X=0) = e^{-1} = 0.3679$

(b) **(5 points)** You pick 5 bases to look at. What is the probability that at least one of those bases is not covered by any reads?

Answer

$Y \sim \mathrm{Binom}(p = 0.3679, n = 5)$
$P(Y \geq 1) = 1 - P(Y = 0) = 1 - C_0^5(0.3679)^0(1 - 0.3679)^5 = 1 - 0.1009 = 0.8991$

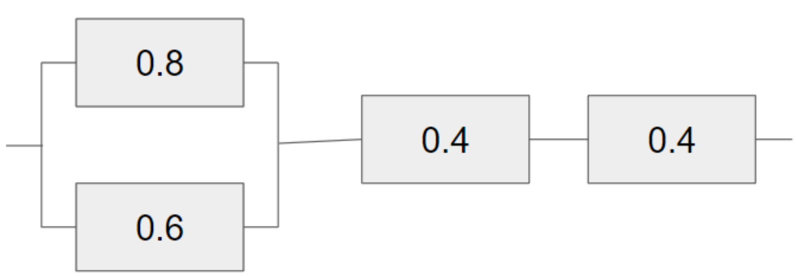
(c) **(5 points)** What is the expected number of bases we have to look at before exactly 20 uncovered bases are identified?
Answer:.
$n_{bases} = 20/0.3679 = 54.36$

4. **(10 points)** Viral loads can sometimes play an important role in the transmissibility of a viruses between hosts. The common logarithm (base 10) of the viral load (virions per milliliter) of SARS-CoV-2 in a randomly selected infected individual is normally distributed with mean μ= 6, and standard dev. σ = 1.5.
   (a) **(5 points)** What is the probability that the expression level measured randomly chosen infected person is between 100,000 and 1,000,000 virions per milliliter?

   Answer: P(-1.333< Z< 0) =P(Z<0)-P(Z<-1.333)= 0.5- 0.2525 = 0.2475

   (b) **(5 points)** The viral loads were measured in 5 individuals. What is the probability that at least 3 individuals' viral loads are within the bounds from part a?
   Answer: Y~Binom(p=0.2475, n=5)
   P(Y>= 3) = P(Y=3) +P(Y=4) + P(Y=5)
   =0.08585 + 0.01412 + 0.00093
   = 0.1009

5. **(10 points)** The following circuit operates if and only if there is a path of functional devices from left to right. The probability that each device functions is as shown. Assume that the probability that a device is functional does not depend on whether or not other devices are functional. What is the probability that the circuit operates?



Answer: P = (1-(1-0.6)*(1-0.8))*0.4*0.7 = 0.2576

6. **(10 points)** You are a part of a group of 20 friends who want to go to a restaurant together. You plan to go there using shuttle vehicles. Since each vehicle can only fit 4 people, you will need multiple vehicles to transport the entire group. Assume the vehicles arrive according to a constant rate (Poisson) process with the mean rate equal to two vehicles per five minutes. Also assume that each vehicle will leave full (4 people per vehicle).
   (a) **(5 points)** What is the probability that a time interval between two vehicles is longer than 3 minutes?

   Answer: lambda=2 message/5 minutes=0.4 messages/minute.
   Exponential distribution P(X>3)=exp(-0.4*3)=exp(-1.2)=0.3012

   |

   (b) **(5 points)** What is the mean time until the entire group is on route to the restaurant?

   Answer: Using Erlang distribution with r=5, lambda=0.4 one gets (5/0.4) minutes =12.5 minutes