Math 321 – Summer 2019

additional practice for chapter 4, central limit theorem

1. If we flip a fair coin many many times, what will the proportion of heads be?

   Solution:

   Due to the law of large numbers, the greater the number of flips, the less likely the proportion of heads is to deviate too far from 50%.

2. Consider that height of a certain demographic category of people has mean 5 foots 9 inches and standard deviation 3 inches. If a group of 30 people of this demographic are randomly selected, what is the probability their mean height is less than 5 foot 8 inches?

   Solution:

   Using the CLT, we know that $\overline{X} \sim N(\mu, \sigma^2/n)$ thus in our random sample of 30, we have $\overline{X} \sim N(5.75, \sigma^2 = 9/\sqrt{30})$. Thus $P(\overline{X} < 5 + 8/12) =$`pnorm(5+8/12,5.75,3/12/sqrt(30))`$\approx 3.4\%$. Note that we have to be careful about units here and be sure to convert inches into feet carefully.

3. An internet server has data request that arrive at rate 1000 requests per second. Model the time between requests by an exponential random variable. Scientists would like to study the average energy consumption of the server and need to know the average length of idle time (time between server requests). In order to estimate the probability that energy consumption exceeds a particular threshold they would like to know the probability that the mean wait time between requests exceeds 1.0015 millisecond for a random sample of 1 million requests.

   Solution:

   We know that the time between requests has mean and standard deviation equal to 0.001 seconds since it is modeled by an exponential with rate parameter $\lambda = 1$ request per millisecond (since 1000 requests per second and 1000 ms = 1 s). So then by the CLT $\overline{X} \sim N(1, 1/\sqrt{(10^6)})$. Thus we calculate $P(\overline{X} > 1.0015) =$`1-pnorm(1.0015,mean=1,sd=1/sqrt(10^6))`$\approx 0.0668072$.

4. Consider a print publisher who has on average 1 error (ink smudge, double printed letter, etc.) every 25 pages. A particular document has 150 pages and 400 copies of it are needed. We are interested in the average number of errors per document.

   (a) Use the Poisson distribution to calculate the probability that the total number of typos in the printed lot (400 copies of the 150 page document) will be greater than 2500. Note that this is identical to there being an average of 6.25 errors or greater per document copy.

   Solution:

   Our base rate of error production is 1/25 (1 error per 25 pages) but we need to scale this up to 400×150 pages, so we will use $\lambda = 1/25 \cdot 400 \cdot 150 = 2400$. We expect to have a total of 2400 errors in the entire printed lot. The actual number of errors in the entire lot will be Poisson with rate parameter $\lambda = 2400$, i.e. that $X \sim \text{Pois}(\lambda = 2400)$.

   In R we calculate this by `1-ppois(2500,lambda=2400)` $= 0.02063767$.

   (b) Use the normal approximation given by the central limit theorem to calculate the probability that the mean number of typos per copy will be greater than 6.25.

   Solution:

   Each copy of the document will have a number of typos that is Poisson distributed with mean $\lambda = 6$. We have 400 copies of this document. The sample average of these 400 copies will be reasonably approximated by a normal with mean 6 ($\mu = \lambda$) and variance 6/400 ($\sigma^2/n = \lambda/n$).

   In R we calculate this as `1-pnorm(6.25,mean=6,sd=sqrt(6/400))` $= 0.02061342$.

   In general, a normal distribution with mean $\mu = \lambda$ and variance $\sigma^2 = \lambda$ will be a good approximation to a Poisson with parameter λ when λ is somewhat large. So we can also calculate this as `1-pnorm(2500,mean=2400,sd=sqrt(2400))` $= 0.02061342$.

5. Consider the distribution give by $x = 6, 10$ and $f(x) = 0.2, 0.8$. Find the sampling distribution of $\overline{X}_2$.

<u>*Solution:*</u>

The possible samples of size 2 are $(6,6), (6,10), (10,6), (10,10)$ and the sample means of these are $6, 8, 8, 10$ respectively.

So the distinct sample mean values are $6, 8, 10$. We tabulate the probability of each.

X_1	X_2	$P(X_1, X_2)$	$\overline{x}$
6	6	$(0.2)^2$	6
6	10	$(0.2)(0.8)$	8
10	6	$(0.8)(0.2)$	8
10	10	$(0.8)^2$	10

The sampling distribution for $\overline{X}$ is given below.

$\overline{x}$	6	8	10
$f_{\overline{X}}(\overline{x})$	0.04	0.32	0.64

Note that this is a probability mass function for random variable $\overline{X} = \frac{1}{2}(X_1 + X_2)$ which is a function of two independence random variables.