1. On any given flight, an airline’s goal is to fill the plane as much as possible, without overbooking. If, on average, 10% of customers cancel their tickets, all independently of each other, what is the probability that a particular flight will be overbooked if the airline sells 320 tickets, for a plane that has maximum capacity 300 people? What is the probability that a plane with maximum capacity 150 people will be overbooked if the airline sells 160 tickets?

Solution: Let $X$ be the number of people who show up for the 300 person plane and $Y$ be the number for the 150 person plane.

$E(X) = 288$, $E(Y) = 144$, $Var(X) = 320(0.1)(0.9) = 28.8$, $Var(Y) = 160(0.1)(0.9) = 14.4$

Then we apply the central limit theorem. Well, really, the DeMoivre-Laplace theorem for the binomial.

\[ P(X > 300) = P \left( \frac{X - 288}{\sqrt{28.8}} > \frac{12}{\sqrt{28.8}} \right) \]
\[ \approx 1 - \Phi \left( \frac{12.5}{\sqrt{28.8}} \right) \]
\[ = 1 - \Phi(2.329) \]

\[ P(Y > 150) = P \left( \frac{Y - 144}{\sqrt{14.4}} > \frac{6}{\sqrt{14.4}} \right) \]
\[ \approx 1 - \Phi \left( \frac{6.5}{\sqrt{14.4}} \right) \]
\[ = 1 - \Phi(1.713) \]

In the last step for each we have used the continuity correction.

2. Let $X_1, \ldots, X_{25}$ be iid $\sim Unif[0,1]$. Let $S$ be their sum

(a) Estimate $P(S \geq 15)$ using Markov’s inequality.

(b) Repeat part (a) using Chebyshev’s inequality.

(c) Repeat part (a) using the central limit theorem.

Solution: $E(S) = 12.5$, $Var(S) = \frac{25}{12}$

(a) $P(S > 15) \leq \frac{12.5}{15}$

(b) $P(|S - 12.5| > 2.5) \leq \frac{12.5/12}{2.5^2}$

(c) $P(S > 15) = P \left( \frac{S - 12.5}{\sqrt{25/12}} > \frac{2.5}{\sqrt{12}} \right) \approx 1 - \Phi(\sqrt{12}/2)$
3. The adult population of Nowhereville consists of 300 males and 196 females. Each male (respectively, female) has a probability of 0.4 (respectively, 0.5) of casting a vote in the local elections, independently of everyone else. Find a good numerical approximation for the probability that more males than females cast a vote.

Solution: Let $M$ be the number of males who vote and $F$ be the number of females who vote. By the central limit theorem, we can approximate $M \sim \mathcal{N}(120, 72)$ and the number for the females is $F \sim \mathcal{N}(98, 49)$.

The Difference is approximately $M - F \sim \mathcal{N}(22, 121)$. So we want the probability that this is more than 0.

$$P(M - F > 0) = 1 - \Phi(-22 / 1) = 1 - \Phi(-2) \approx 0.95$$

4. Bo assumes that $X$, the height in meters of any Canadian selected by an equally likely choice among all Canadians, is a random variable with $E(X) = h$. Because Bo is sure that no Canadian is taller than 3 meters, he decides to use 1.5 meters as a conservative value for the standard deviation of $X$.

To estimate $h$, Bo uses the average of the heights of $n$ Canadians he selects at random.

(a) In terms of $h$ and Bo’s 1.5 meter bound for the standard deviation of $X$ determine the expectation and standard deviation of $H$.

(b) Find as small a value of $n$ as possible such that the standard deviation of Bo’s estimator is guaranteed to be less than 0.01 meters.

(c) Bo would like to be 99% sure that his estimate is within 5 centimeters of the true average height of Canadians. Using the Chebyshev inequality, calculate the minimum value of $n$ that will achieve this objective.

Solution: Let $X_1, \ldots, X_n$ be the heights of the $n$ people he samples and $H = \frac{\sum_{i=1}^{n} X_i}{n}$.

(a) $E(H) = h$, $Var(H) = \frac{1.5^2}{n}$

(b) We want $\frac{1.5}{n} \leq 0.01$. This implies $n \geq 150$.

(c) By Chebyshev, $P(|H - h| \geq 0.05) \leq \frac{2.25/n}{(1/20)^2} = 900/n$. We want $\frac{900}{n} \leq 0.01$ so we have $n \geq 90000$.

5. (a) Find the PDF and CDF of $U_1 + U_2$ where they are iid $\sim Unif[0, 1]$ random variables. It will be done piecewise.

(b) Find the PDF and CDF of the product $U_1U_2$. Hint, what is $-\log(U_1)$?

Solution:

(a) We want to calculate $\int_{-\infty}^{\infty} f_{U_1}(x) f_{U_2}(z - x) dx$. Notice that each of these densities is 1 when its argument is between 0 and 1, and 0 otherwise. So, the product of the densities is 1 when both are 1 and 0 otherwise. The first term in the integrand tells us we need $x \geq 0$. The second is more complicated. It says that if $0 \leq z \leq 1$ then $x$ is between 0 and $x$ (because $z - x < 0$ if $x > z$). If $z > 1$, however, $z - x$ can still be less than 1 if $z - 1 \leq x \leq z$. However, since $x \leq 1$ as well, we integrate from $z - 1$ to 1 in this case. Also, note that $z \leq 2$. Finally we get

$$f_{U_1+U_2}(z) = \int_{-\infty}^{\infty} dx = \begin{cases} \int_{0}^{z} dx & 0 \leq z \leq 1 \\ \int_{z-1}^{1} dx & 1 < z \leq 2 \\ \end{cases} = \begin{cases} z & 0 \leq z \leq 1 \\ 2 - z & 1 < z \leq 2 \\ \end{cases}$$
To find the CDF, we just integrate. Note that we have to do it piecewise.

\[ F_{U_1+U_2}(z) = \begin{cases} 
  z^2/2, & 0 \leq z \leq 1 \\
 2z - z^2/2 - 1 & \end{cases} \]

(b) Notice that \( \log(U_1) \sim Expo(1) \) and similarly for \( U_2 \). Letting \( U = U_1U_2 \) we have \(-\log(U) = -\log(U_1) - \log(U_2)\). Knowing what we know about the sum of exponentials, we have

\[ F_{-\log(U)}(z) = \int_0^z x e^{-x} \, dx = 1 - e^{-z}(z+1) = P(-\log(U) \leq z) = P(U \geq e^{-z}) = 1 - P(U \leq e^{-z}) \]

Letting \( z = -\log(w) \) we have

\[ 1 - w(1 - \log(w)) = 1 - P(U \leq w) \]

and thus \( F_U(w) = w(1 - \log(w)) \) for \( 0 \leq w \leq 1 \). To find the CDF just differentiate and we get \( f_U(z) = -\log(w) \)

6. Let \( X \) be a random variable with MGF \( E(e^{tX}) = a + be^{2s} + ce^{4s} \) and suppose we know \( E(X) = 3, Var(X) = 2 \). Find \( a, b, c \) and the PMF of \( X \).

Solution: Differentiating and setting \( t = 0 \) to find the moments tells us that \( E(X) = 2b + 4c = 3 \)

and \( E(X^2) = 4b + 16c = 11 \). From this we get \( b = 1/4, c = 5/8 \). Since \( a + b + c = 1 \), \( a = 1/8 \)

7. Suppose \( E(e^{tX}) = \frac{6-3t}{2(1-t)(3-t)} \). What is the PDF of \( X \)?

Solution: The trick here is to notice that this can be written as \( \frac{1}{3-2t} + \frac{3}{2(1-t)} \). When we do this, we see that this is a mixture of exponential distributions, so the pdf is \( \frac{1}{5}e^{-3x} + \frac{3}{4}e^{-x} \).

8. I flip a fair coin, and if it shows heads I sample from an \( Expo(1) \) random variable and if it is tails from a \( N(0,1) \). Call this \( X \). Compute the MGF of \( X \), \( E(X) \) and \( Var(X) \)

Solution: We use the law of total expectation

\[ E(e^{tX}) = E(e^{tX} | H) P(H) + E(e^{tX} | T) P(T) = \frac{1}{2} \left( \frac{1}{1-t} + e^{-t^2/2} \right) \]