# CSE 312 HW5

## Rohan Mukherjee

## May 16, 2023

1. Note: throughout this problem, I put $\approx$ signs since the value of $\Phi$'s are not exact (and, I might plug in a rounded value into $\phi$).

   Since $X$ has mean 30 and standard deviation 15, we know that $\frac{X-30}{15}$ is a standardized normally distributed variable. Then,

$$\begin{aligned}
\mathbb{P}(25 < X < 65) = \mathbb{P}(25 \leq X \leq 65) &= \mathbb{P}\left(\frac{25-30}{15} \leq \frac{X-30}{15} \leq \frac{65-30}{15}\right) \\
&\approx \mathbb{P}\left(-0.33 \leq \frac{X-30}{15} \leq 2.33\right) \\
&= \Phi(2.33) - \Phi(-0.33) \\
&= \Phi(2.33) - (1 - \Phi(0.33)) \\
&= 0.9901 - (1 - 0.6293) = 0.6194
\end{aligned}$$

   The approximation is here since one can only approximate the CDF of the normal distribution.

   (b) Let $X$ be the r.v. representing the weight of a baby. The problem tells us that it's mean is 3400 and standard deviation is 500, so $\frac{X-3400}{500}$ is a standard normal variable. We are looking for

$$\mathbb{P}(X > 4100) = 1 - \mathbb{P}(X \leq 4100) = 1 - \mathbb{P}\left(\frac{X-3400}{500} < 1.4\right) = 1 - \Phi(1.4) = 0.08076$$

   So about $8076/100,000$ of the babies will weigh more than 4100 grams.

   (c) Big fan of this question! If $X$'s rap beats $Y$'s rap, then $X > Y$, so motivated by this let $Z = X - Y$. We are therefore looking for $\mathbb{P}(Z > 0)$. We need to find the standard deviation and variance of $Z$ so that we can normalize it. $\mathbb{E}[X-Y] = \mathbb{E}[X] - \mathbb{E}[Y] = 5 - 4 = 1$. Also, $\text{Var}(X-Y) = \text{Var}(X) + (-1)^2\text{Var}(Y) = 2 + 5 = 7$ (Recall, the rap quality is independent). We conclude that

$$\mathbb{P}(Z > 0) = 1 - \mathbb{P}(Z \leq 0) = 1 - \mathbb{P}\left(\frac{Z-1}{7} \leq -\frac{1}{7}\right)$$
$$\approx 1 - \Phi(-0.14) = 1 - (1 - \Phi(0.14)) = \Phi(0.14) = 0.55567$$

   Quite a nice answer!

2. (a) Let $X_i$ be the amount of money you have to pay back friend $i \in \{1, \ldots, 20\}$, and let $X = \sum_{i=1}^{20} X_i$. We see that $\mathbb{E}[X] = 20 \cdot 127 = 2540$ and that $\text{Var}(X) = \sum_{i=1}^{20} \text{Var}(X_i) = 20 \cdot 35 = 700$. We want to find $k$ so that

$$\mathbb{P}(X \leq k) \geq 0.99$$

And then we should take out a mortgage of size $k$. This is saying that the money I take out ($k$) is at least the money I owe to my friends 99% of the time. First standardizing, we see that

$$\mathbb{P}(X \leq k) = \mathbb{P}\left(\frac{X - 2540}{700} \leq \frac{k - 2540}{700}\right)$$

Letting $u = \frac{k-2540}{700}$ temporarily, we are looking for $u$ so that (Since we normalized our variable), $\Phi(u) \geq 0.99$. Going to the Z-table, we see that the smallest such $u$ is $u = 2.33$. Solving back for $k$ gives $k = 4171$. So we would need to take out a mortgage of $4171, which is quite a bit of money.

(b) Let $Z_i$ be the net return of the $i$th cryptocurrency ($i \in \{1, \ldots, 8\}$), and let $Z = \sum_{i=1}^{8} Z_i$. We see that $\mathbb{E}[Z] = 8 \cdot \mathbb{E}[Z_i] = 8 \cdot (38,000 \cdot p - 12,000 \cdot (1-p)) = 400,000p - 96,000$. Also, one sees that

$$\text{Var}(Z_1) = \mathbb{E}[Z_1^2] - \mathbb{E}[Z_1]^2 = (38,000^2 \cdot p + 12,000^2 \cdot (1-p)) - (38,000 \cdot p - 12,000 \cdot (1-p))^2$$
$$= 2,500,000,000p(1-p)$$

And hence, since the $Z_i$'s are independent, $\text{Var}(Z) = 20,000,000,000p(1-p)$ (Again, 8 times the last quantity).

We wish to find $p$ so that

$$\mathbb{P}(Z > 0) \geq 0.99$$

Since $Z$ is the sum of discrete variables, it is also discrete, and we need to use continuity correction. Since $Z$ cannot take on the values of $[-1, -12,000 + 1]$, we shall use $(-12,000 + 1 - 1)/2 = -6,000$ as the new lower bound (right in the middle). We consider

$$\mathbb{P}(Z > -6000) \geq 0.99$$

Which is equivalent to saying $\mathbb{P}(Z \leq -6000) \leq 0.01$ by complimentary counting. Notice that

$$\mathbb{P}(Z \leq -6000) = \mathbb{P}\left(\frac{Z - (400,000p - 96,000)}{\sqrt{20,000,000,000p(1-p)}} \leq \frac{-6,000 - (400,000p - 96,000)}{\sqrt{20,000,000,000p(1-p)}}\right)$$

The RHS equals $\Phi\left(\frac{90,000-400,000p}{\sqrt{20,000,000,000p(1-p)}}\right)$ since we standardized the r.v.. We shall now let $u = \frac{90,000-400,000p}{\sqrt{20,000,000,000p(1-p)}}$. We want $\mathbb{P}(u) \leq 0.01$, but 0.01 is not on the table, so instead we shall notice that $\mathbb{P}(u) = 1 - \mathbb{P}(-u)$, and then we see that we want $\mathbb{P}(-u) \geq 0.99$. Once again the first time this happens is $-u = 2.33$. Solving for $p$ gives us $p \approx 0.62402$, which seems like it's in the right ball park.

2

(c) Let $X_i$ be 1 if the $i$th person was hired and 0 otherwise $i \in \{1, \ldots n\}$. Clearly $X = \sum_{i=1}^{n} X_i$ is the total number of people hired among the $n$. We want to find $n$ so that $\mathbb{P}(X \geq 312) \geq 0.95$. First, notice that $\mathbb{E}[X_i] = 0.75$ as was given in the problem statement (expectation of an indicator random variable). It follows that $\mathbb{E}[X] = 0.75n$ by linearity of expectation. Also, $\text{Var}(X_i) = \mathbb{E}[X_i^2] - \mathbb{E}[X_i]^2 = \mathbb{E}[X_i] - \mathbb{E}[X_i]^2 = 0.75 - 0.75^2 = 0.1875$, and since the $X_i$'s are independent, $\text{Var}(X) = 0.1875n$ (pull the sum out of the variance). Unlike last time, $X$ can actually take on the value of 311, so this time our lower bound must be 311.5 for the continuity correction (since the number of people you hire is an integer). Notice now that $\frac{X - 0.75n}{\sqrt{0.1875n}}$ is a standard normal, and

$$\mathbb{P}(X \geq 312) = \mathbb{P}(X \geq 311.5) = \mathbb{P}\left(\frac{X - 0.75n}{\sqrt{0.1875n}} \geq \frac{311.5 - 0.75n}{\sqrt{0.1875n}}\right)$$
$$= 1 - \mathbb{P}\left(\frac{X - 0.75n}{\sqrt{0.1875n}} \leq \frac{311.5 - 0.75n}{\sqrt{0.1875n}}\right) \approx 1 - \Phi\left(\frac{311.5 - 0.75n}{\sqrt{0.1875n}}\right)$$

We want the RHS to be $\geq 0.95$, which is equivalent to saying $\Phi\left(\frac{311.5 - 0.75n}{\sqrt{0.1875n}}\right) \leq 0.05$. Letting $u = \frac{311.5 - 0.75n}{\sqrt{0.1875n}}$ temporarily, we see that we want $\Phi(u) \leq 0.05$. However, this value is not on the table, so we need to use the same trick as in part (a). This is equivalent to saying $\Phi(-u) \geq 0.95$, which gives us $-u = 1.65$. Solving back for $n$ gives $n \approx 435.207$, but since we can only meet with an integer amount of people, we must take $n = \lceil 435.207 \rceil = 436$, which seems in the right ball park.

3. First, I appreciate you guys giving a really easy question after a really long one. In last quarter, I was in Math 335, and my professor also gave a little talk about the problem with polling. He carefully proved a bunch of facts about Monte-Carlo simulations, but most importantly that (independent of dimension, which is crazy), it will have it's standard deviation relating to $C \cdot \frac{1}{\sqrt{n}}$ where $n$ is the number of people sampled (and $C > 0$ is some constant). When polls survey 300 people, they can be off by 6%, which is a lot where every percent matters (e.g., getting 49% vs 51% means literally nothing in this case). Also, most of the polls are not truly random–I believe the only people willing to take a political poll are the people that have strong opinions about the topic at hand (e.g., if you get a phone call asking you to take a political poll, most people in the middle would hang up right away. However, if you have a strong opinion you will not hang up and instead give said strong opinion). Hence, the votes are skewed by those at the far ends of the political spectrum.

4. (a) When $\lambda \leq 0$, the integral explodes off to $\infty$ (unless $c = 0$, in which case the integral is 0 and in particular not 1). This is because

$$c \int_{\mathbb{R}} e^{-2\lambda|x|} dx = 2c \int_0^\infty e^{-2\lambda x} dx = -\frac{c}{\lambda} e^{-2\lambda x} \Big|_0^\infty$$

$$= -\frac{c}{\lambda} \left( \lim_{x \to \infty} e^{-2\lambda x} - 1 \right) \to \pm\infty$$

Plus or minus depending on the sign of $c$. In any case it isn't 1, and can't be made into 1 given any choice of $c$ (either $c$ is not zero or zero, we already handled the case where $c$ is zero. If it not 0, $\infty \cdot c$ is certainly still $\infty$).

(b) Assuming that $\lambda > 0$, we see that $e^{-2\lambda x} \overset{x \to \infty}{\to} 0$, so the integral (which I found above) equals $\frac{c}{\lambda}$. We want this to equal 1, so $c = \lambda$.

(c) By definition, the mean (expected value) of $X$ is

$$\lambda \int_{\mathbb{R}} x e^{-2\lambda|x|} dx = 0$$

This makes sense since our function is odd and this integral clearly converges. Next, the variance is

$$\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \lambda \int_{\mathbb{R}} x^2 e^{-2\lambda|x|} dx - 0 = \frac{1}{2\lambda^2}$$

Which is nice.

(d) First, if $x = 0$, then we are looking for

$$\mathbb{P}(X \geq 0) = 1 - \mathbb{P}(X \leq 0) = 1 - \int_{-\infty}^0 \lambda e^{-2\lambda|t|} dt = 1 - \lambda \int_0^\infty e^{-2\lambda t} dt$$

$$= 1 + \frac{1}{2} e^{-2\lambda t} \Big|_0^\infty = \frac{1}{2}$$

Which is very intuitive. One notices this also tells us (after rearranging) that

$$\int_{-\infty}^0 \lambda e^{-2\lambda|t|} dt = \frac{1}{2}$$

If $x > 0$, we are looking for

$$\mathbb{P}(X \geq x) = 1 - \mathbb{P}(X \leq x) = 1 - \int_{-\infty}^x \lambda e^{-2\lambda|t|} dt$$

$$= 1 - \left( \int_{-\infty}^0 \lambda e^{-2\lambda|t|} dt + \int_0^x \lambda e^{-2\lambda t} dt \right)$$

$$= 1 - \left( \frac{1}{2} - \frac{1}{2} e^{-2\lambda t} \Big|_0^x \right) = 1 - \left( \frac{1}{2} - \frac{1}{2} (e^{-2\lambda x} - 1) \right) = \frac{1}{2} e^{-2\lambda x}$$

5

Finally, for $x < 0$, we are looking for

$$\mathbb{P}(X \geq x) = 1 - \mathbb{P}(X \leq x) = 1 - \int_{-\infty}^{x} \lambda e^{-2\lambda|t|} dt = 1 - \left( \int_{-\infty}^{x} \lambda e^{-2\lambda|t|} dt + \int_{x}^{0} \lambda e^{-2\lambda|t|} dt - \int_{x}^{0} \lambda e^{-2\lambda|t|} dt \right)$$

$$= 1 - \left( \int_{-\infty}^{0} \lambda e^{-2\lambda|t|} dt - \int_{x}^{0} \lambda e^{-2\lambda|t|} dt \right) = 1 - \left( \frac{1}{2} - \int_{0}^{-x} \lambda e^{-2\lambda t} dt \right) \quad \text{(By symmetry)}$$

Now, $-x$ is positive, so we can look to our answer from last time to get this integral.
I colored everything, so

$$\int_{0}^{-x} \lambda e^{-2\lambda t} dt = -\frac{1}{2}\left( e^{-2\lambda(-x)} - 1 \right) = \frac{1}{2} - \frac{1}{2} e^{2\lambda x}$$

Our final answer for $\mathbb{P}(X \geq x)$ is then $\frac{1}{2} + \int_{0}^{-x} \lambda e^{-2\lambda t} dt = 1 - \frac{1}{2} e^{2\lambda x}$.

5. We need to find $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$. The problem gives us that $\mathbb{E}[X] = \frac{1}{m+1}$, so we are left to evaluate $\mathbb{E}[X^2]$. Noting that we need only integrate our density function against $x^2$ in $[0, 1]$ (the minimum of things in $[0, 1]$ will also lie in $[0, 1]$) we only have to differentiate the piece of the CDF that is defined for $[0, 1]$. Looking at the textbook, the CDF for $x \in [0, 1]$ is $1 - (1 - x)^m$. Differentiating this gives $f_X(x) = m(1 - x)^{m-1}$ for $x \in [0, 1]$. Now,

$$\mathbb{E}[X^2] = \int_0^1 x^2 m(1-x)^{m-1} dx = m \int_0^1 x^2 (1-x)^{m-1} dx$$

We evaluate this integral by integration by parts:

| $D$ | | $I$ |
|---|---|---|
| $x^2$ | $+$ | $(1-x)^{m-1}$ |
| $2x$ | $-$ | $\frac{-1}{m}(1-x)^m$ |
| $2$ | $+$ | $\frac{1}{m(m+1)}(1-x)^{m+1}$ |
| $0$ | | $\frac{-1}{m(m+1)(m+2)}(1-x)^{m+2}$ |

So,

$$m \int_0^1 x^2 (1-x)^{m-1} dx = m \left( \frac{-x^2}{m}(1-x)^m \Big|_0^1 - \frac{2x}{m(m+1)}(1-x)^{m+1} \Big|_0^1 - \frac{2}{m(m+1)(m+2)}(1-x)^{m+2} \Big|_0^1 \right)$$

$$= m \left( 0 - 0 - (0 - 0) - \left( 0 - \frac{2}{m(m+1)(m+2)} \right) \right) = \frac{2}{(m+1)(m+2)}$$

We conclude that

$$\text{Var}(X) = \frac{2}{(m+1)(m+2)} - \frac{1}{(m+1)^2}$$