Math 302.102 Fall 2010
Solutions to Assignment #8

1. Since X has a binomial distribution with parameters $n = 100$ and $p = 4/7$, we can use the central limit theorem to approximate $\mathbf{P}\{X \geq 50\}$. That is, we know that if $X \sim \text{Bin}(n, p)$, then

$$\frac{X - np}{\sqrt{np(1-p)}}$$

has an approximately normal $\mathcal{N}(0, 1)$ distribution as long as both np and $n(1-p)$ are not too small.

Solution 1. Using the continuity correction, we find

$$\begin{aligned}\mathbf{P}\{X \geq 50\} = \mathbf{P}\{X > 49.5\} &= \mathbf{P}\left\{\frac{X - 100 \cdot \frac{4}{7}}{\sqrt{100 \cdot \frac{4}{7} \cdot \frac{3}{7}}} > \frac{49.5 - 100 \cdot \frac{4}{7}}{\sqrt{100 \cdot \frac{4}{7} \cdot \frac{3}{7}}}\right\} \\ &\approx \mathbf{P}\{Z > -1.544\}\end{aligned}$$

where $Z \sim \mathcal{N}(0, 1)$. Using a table of normal probabilities, we find $\mathbf{P}\{Z > -1.544\} \doteq 0.9382$. Thus,

$$\mathbf{P}\{X \geq 50\} \approx 0.9382.$$

Solution 2. Observe that the largest X could be is 100. Thus, $\mathbf{P}\{X \geq 50\}$ is really equivalent to $\mathbf{P}\{50 \leq X \leq 100\}$, and so using the continuity correction, we find

$$\begin{aligned}\mathbf{P}\{X \geq 50\} = \mathbf{P}\{50 \leq X \leq 100\} &= \mathbf{P}\{49.5 < X < 100.5\} \\ &= \mathbf{P}\left\{\frac{49.5 - 100 \cdot \frac{4}{7}}{\sqrt{100 \cdot \frac{4}{7} \cdot \frac{3}{7}}} < \frac{X - 100 \cdot \frac{4}{7}}{\sqrt{100 \cdot \frac{4}{7} \cdot \frac{3}{7}}} < \frac{100.5 - 100 \cdot \frac{4}{7}}{\sqrt{100 \cdot \frac{4}{7} \cdot \frac{3}{7}}}\right\} \\ &\approx \mathbf{P}\{-1.544 < Z < 8.761\} \\ &= \mathbf{P}\{Z < 8.761\} - \mathbf{P}\{Z < -1.544\}\end{aligned}$$

where $Z \sim \mathcal{N}(0, 1)$. Using a table of normal probabilities, we find $\mathbf{P}\{Z < 8.761\} = 1$ and $\mathbf{P}\{Z < -1.544\} \doteq 0.0618$. Thus,

$$\mathbf{P}\{X \geq 50\} \approx 1 - 0.0618 = 0.9382.$$

Solution 3. Without using the continuity correction, we find

$$\begin{aligned}\mathbf{P}\{X \geq 50\} &= \mathbf{P}\left\{\frac{X - 100 \cdot \frac{4}{7}}{\sqrt{100 \cdot \frac{4}{7} \cdot \frac{3}{7}}} \geq \frac{50 - 100 \cdot \frac{4}{7}}{\sqrt{100 \cdot \frac{4}{7} \cdot \frac{3}{7}}}\right\} \\ &\approx \mathbf{P}\{Z > -1.443\}\end{aligned}$$

where $Z \sim \mathcal{N}(0, 1)$. Using a table of normal probabilities, we find $\mathbf{P}\{Z > -1.443\} \doteq 0.9251$. Thus,

$$\mathbf{P}\{X \geq 50\} \approx 0.9251.$$

Remark. The exact probability can be determined using a computer. It is 0.9381732. Note that the approximations provided by Solutions 1 and 2 are remarkably accurate.

2. Let X_j denote the outcome of the jth game so that $Y_j = X_j - 4$ denotes the net winnings on the jth game. The total net winnings are therefore

$$Y_1 + \cdots + Y_{100} = X_1 + \cdots + X_{100} - 400.$$

Let

$$\overline{Y} = \frac{Y_1 + \cdots + Y_{100}}{100} \quad \text{and} \quad \overline{X} = \frac{X_1 + \cdots + X_{100}}{100}$$

so that

$$\overline{Y} = \frac{Y_1 + \cdots + Y_{100}}{100} = \frac{X_1 + \cdots + X_{100} - 400}{100} = \frac{X_1 + \cdots + X_{100}}{100} - \frac{400}{100} = \overline{X} - 4.$$

It is equivalent to work with either $\overline{Y}$ or $\overline{X}$. The basic idea is that the central limit theorem tells us that both

$$\frac{\overline{Y} - \mathbb{E}(\overline{Y})}{\sqrt{\text{var}(\overline{Y})}} \quad \text{and} \quad \frac{\overline{X} - \mathbb{E}(\overline{X})}{\sqrt{\text{var}(\overline{X})}}$$

have an approximately normal $\mathcal{N}(0,1)$ distribution. Since $X_1, X_2, \ldots, X_{100}$ are iid, their common mean is

$$\begin{aligned}
&\mathbb{E}(X_1) \\
&= 1 \cdot \mathbf{P}\{X_1 = 1\} + 2 \cdot \mathbf{P}\{X_1 = 2\} + 3 \cdot \mathbf{P}\{X_1 = 3\} + 4 \cdot \mathbf{P}\{X_1 = 4\} + 5 \cdot \mathbf{P}\{X_1 = 5\} + 6 \cdot \mathbf{P}\{X_1 = 6\} \\
&= \frac{1}{6}(1+2+3+4+5+6) = \frac{21}{6} = 3.5
\end{aligned}$$

and, similarly,

$$\mathbb{E}(X_1^2) = \frac{1}{6}(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) = \frac{91}{6}$$

so that their common variance is

$$\text{var}(X_1) = \mathbb{E}(X_1^2) - [\mathbb{E}(X_1)]^2 = \frac{91}{6} - \left(\frac{21}{6}\right)^2 = \frac{105}{36} = \frac{35}{12}.$$

Thus,

$$\mathbb{E}(\overline{X}) = \mathbb{E}(X_1) = 3.5 \quad \text{and} \quad \text{var}(\overline{X}) = \frac{\text{var}(X_1)}{100} = \frac{35}{1200} = \frac{7}{240}.$$

The fact that $\overline{Y} = \overline{X} - 4$ implies that

$$\mathbb{E}(\overline{Y}) = \mathbb{E}(\overline{X} - 4) = \mathbb{E}(\overline{X}) - 4 = 3.5 - 4 = -0.5 \quad \text{and} \quad \text{var}(\overline{Y}) = \text{var}(\overline{X} - 4) = \text{var}(\overline{X}) = \frac{7}{240}.$$

2. (a) The probability that you will lose exactly \$50 is

$$\mathbf{P}\{Y_1 + \cdots + Y_{100} = -50\}.$$

Using the continuity correction we find

$$\begin{aligned}
\mathbf{P}\{Y_1 + \cdots + Y_{100} = -50\} &= \mathbf{P}\{-50.5 < Y_1 + \cdots + Y_{100} < -49.5\} \\
&= \mathbf{P}\left\{-\frac{50.5}{100} < \frac{Y_1 + \cdots + Y_{100}}{100} < -\frac{49.5}{100}\right\} \\
&= \mathbf{P}\left\{-\frac{50.5}{100} < \overline{Y} < -\frac{49.5}{100}\right\}.
\end{aligned}$$

Normalizing gives

$$\mathbf{P}\left\{-\frac{50.5}{100} < \overline{Y} < -\frac{49.5}{100}\right\} = \mathbf{P}\left\{\frac{-\frac{50.5}{100}-(-0.5)}{\sqrt{7/240}} < \frac{\overline{Y}-(-0.5)}{\sqrt{7/240}} < \frac{-\frac{49.5}{100}-(-0.5)}{\sqrt{7/240}}\right\}$$
$$\approx \mathbf{P}\left\{-0.029 < Z < 0.029\right\}$$

where $Z \sim \mathcal{N}(0,1)$. Using a table of normal probabilities, we find

$$\mathbf{P}\left\{-0.029 < Z < 0.029\right\} \doteq 0.5120 - 0.4880 = 0.024$$

so that the probability you will lose exactly \$50 is approximately 0.024.

2. (b) The probability that you will win is

$$\mathbf{P}\left\{Y_1 + \cdots + Y_{100} > 0\right\} = 1 - \mathbf{P}\left\{Y_1 + \cdots + Y_{100} \le 0\right\}.$$

Using the continuity correction, we find

$$\mathbf{P}\left\{Y_1 + \cdots + Y_{100} \le 0\right\} = \mathbf{P}\left\{Y_1 + \cdots + Y_{100} \le 0.5\right\} = \mathbf{P}\left\{\overline{Y} \le \frac{0.5}{100}\right\},$$

and normalizing gives

$$\mathbf{P}\left\{\frac{\overline{Y}-(-0.5)}{\sqrt{7/240}} \le \frac{\frac{0.5}{100}-(-0.5)}{\sqrt{7/240}}\right\} \approx \mathbf{P}\left\{Z \le 2.898\right\}$$

where $Z \sim \mathcal{N}(0,1)$. Using a table of normal probabilities, we find

$$\mathbf{P}\left\{Z \le 2.898\right\} \doteq 0.9981$$

so that the probability you will win is approximately $1 - 0.9981 = 0.0019$.

3. (a) Note that if you roll a 1 so that $X = 1$, then no matter what my first roll is, I will always obtain a number that is greater than or equal to yours. Hence, by the definition of conditional probability

$$\mathbf{P}\left\{X = 1, Y = 1\right\} = \mathbf{P}\left\{Y = 1 \,|\, X = 1\right\} \cdot \mathbf{P}\left\{X = 1\right\} = 1 \cdot \frac{1}{6} = \frac{1}{6}.$$

Furthermore, since $\mathbf{P}\left\{Y = 1 \,|\, X = 1\right\} = 1$ it follows that $\mathbf{P}\left\{Y = k \,|\, X = 1\right\} = 0$ for any $k = 2, 3, \ldots$ so that

$$\mathbf{P}\left\{X = 1, Y = k\right\} = 0$$

if $k = 2, 3, \ldots$.

3. (b) Note that if you roll a 2 so that $X = 2$, then there are 5 ways that I can obtain a number that is greater than or equal to yours on any given roll. Thus, the only way that I will re-roll is if I roll a 1 which happens with probability $1/6$. This means that I will need exactly k rolls if my first $k-1$ rolls yield a 1 and my kth roll yields anything other than a 1. This means that

$$\mathbf{P}\left\{Y = k \,|\, X = 2\right\} = \left(\frac{1}{6}\right)^{k-1} \cdot \frac{5}{6}$$

so that by the definition of conditional probability

$$\mathbf{P}\left\{X=2, Y=k\right\} = \mathbf{P}\left\{Y=k \,|\, X=2\right\} \cdot \mathbf{P}\left\{X=2\right\} = \left(\frac{1}{6}\right)^{k-1} \cdot \frac{5}{6} \cdot \frac{1}{6}$$

for $k = 1, 2, 3, \ldots$.

3. (c) Using exactly the same reasoning in the previous part implies that if $k = 1, 2, 3, \ldots$, then

$$\mathbf{P}\left\{X=3, Y=k\right\} = \left(\frac{2}{6}\right)^{k-1} \cdot \frac{4}{6} \cdot \frac{1}{6},$$

$$\mathbf{P}\left\{X=4, Y=k\right\} = \left(\frac{3}{6}\right)^{k-1} \cdot \frac{3}{6} \cdot \frac{1}{6},$$

$$\mathbf{P}\left\{X=5, Y=k\right\} = \left(\frac{4}{6}\right)^{k-1} \cdot \frac{2}{6} \cdot \frac{1}{6},$$

and

$$\mathbf{P}\left\{X=6, Y=k\right\} = \left(\frac{5}{6}\right)^{k-1} \cdot \frac{1}{6} \cdot \frac{1}{6}.$$

3. (d) Using the hint given in the problem, we want to compute

$$\mathbb{E}(Y) = \sum_{i=1}^{6} \sum_{k=1}^{\infty} k\mathbf{P}\left\{X=i, Y=k\right\}.$$

Now, let's consider each of $i = 1, 2, 3, 4, 5, 6$ in turn. For $i = 1$, since the only non-zero value of $\mathbf{P}\left\{X=1, Y=k\right\}$ is when $k = 1$, we have

$$\sum_{k=1}^{\infty} k\mathbf{P}\left\{X=1, Y=k\right\} = 1 \cdot \mathbf{P}\left\{X=1, Y=1\right\} + \sum_{k=2}^{\infty} k\mathbf{P}\left\{X=1, Y=k\right\} = \frac{1}{6} + 0 = \frac{1}{6}.$$

For $i = 2$, we have

$$\sum_{k=1}^{\infty} k\mathbf{P}\left\{X=2, Y=k\right\} = \sum_{k=1}^{\infty} k\left(\frac{1}{6}\right)^{k-1} \cdot \frac{5}{6} \cdot \frac{1}{6} = \frac{5}{6} \cdot \frac{1}{6} \sum_{k=1}^{\infty} k\left(\frac{1}{6}\right)^{k-1} = \frac{5}{6} \cdot \frac{1}{6} \cdot \frac{1}{(1-1/6)^2} = \frac{1}{5}.$$

Similarly, for $i = 3, 4, 5, 6$, we find

$$\sum_{k=1}^{\infty} k\mathbf{P}\left\{X=3, Y=k\right\} = \sum_{k=1}^{\infty} k\left(\frac{2}{6}\right)^{k-1} \cdot \frac{4}{6} \cdot \frac{1}{6} = \frac{4}{6} \cdot \frac{1}{6} \cdot \frac{1}{(1-2/6)^2} = \frac{1}{4},$$

$$\sum_{k=1}^{\infty} k\mathbf{P}\left\{X=4, Y=k\right\} = \sum_{k=1}^{\infty} k\left(\frac{3}{6}\right)^{k-1} \cdot \frac{3}{6} \cdot \frac{1}{6} = \frac{3}{6} \cdot \frac{1}{6} \cdot \frac{1}{(1-3/6)^2} = \frac{1}{3},$$

$$\sum_{k=1}^{\infty} k\mathbf{P}\left\{X=5, Y=k\right\} = \sum_{k=1}^{\infty} k\left(\frac{4}{6}\right)^{k-1} \cdot \frac{2}{6} \cdot \frac{1}{6} = \frac{2}{6} \cdot \frac{1}{6} \cdot \frac{1}{(1-4/6)^2} = \frac{1}{2},$$

and

$$\sum_{k=1}^{\infty} k\mathbf{P}\left\{X=6, Y=k\right\} = \sum_{k=1}^{\infty} k\left(\frac{5}{6}\right)^{k-1} \cdot \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{(1-5/6)^2} = 1.$$

Hence, we can now sum over i to conclude

$$\mathbb{E}(Y) = \sum_{i=1}^{6}\sum_{k=1}^{\infty} k\mathbf{P}\left\{X = i, Y = k\right\} = \frac{1}{6} + \frac{1}{5} + \frac{1}{4} + \frac{1}{3} + \frac{1}{2} + 1 = \frac{49}{20}.$$

4. Since X and Y are independent, we know that

$$\mathbb{E}(X^2 e^Y) = \mathbb{E}(X^2)\mathbb{E}(e^Y).$$

Since $X \sim \mathcal{N}(3, 2)$, we know that $\mathbb{E}(X^2) = \text{var}(X) + [\mathbb{E}(X)]^2 = 2 + 3^2 = 11$. Since $Y \sim \text{Exp}(2)$, we know that

$$\mathbb{E}(e^Y) = \int_0^\infty e^y \cdot 2e^{-2y}\,\mathrm{d}y = 2\int_0^\infty e^{-y}\,\mathrm{d}y = 2.$$

Thus,

$$\mathbb{E}(X^2 e^Y) = 11 \cdot 2 = 22.$$

5. (a) To begin, observe that $\{(x, y) : 0 < x < y < 1\}$ describes the triangle in the first quadrant above the line $y = x$ and below the line $y = 1$. The area of this triangle is therefore

$$\frac{1}{2} \cdot \text{base} \cdot \text{height} = \frac{1}{2} \cdot 1 \cdot 1 = \frac{1}{2}.$$

Since (X, Y) is uniformly distributed, its density must be inversely proportional to the area. That is,

$$f(x, y) = \begin{cases} 2, & 0 < x < y < 1, \\ 0, & \text{otherwise.} \end{cases}$$

5. (b) To compute $\mathbf{P}\left\{X < Y^2\right\}$ we need to integrate the density function over a suitable region. That is,

$$\mathbf{P}\left\{X < Y^2\right\} = \iint\limits_{\{0<x<y^2<1\}} f(x, y)\,\mathrm{d}x\,\mathrm{d}y = \int_0^1 \int_0^{y^2} 2\,\mathrm{d}x\,\mathrm{d}y = \int_0^1 2y^2\,\mathrm{d}y = \frac{2}{3}.$$

5. (c) By definition,

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\,\mathrm{d}y = \int_x^1 2\,\mathrm{d}y = 2(1 - x)$$

provided that $0 < x < 1$.

5. (d) We being by computing

$$\mathbb{E}(X) = \int_0^1 x \cdot 2(1 - x)\,\mathrm{d}x = \int_0^1 (2x - 2x^2)\,\mathrm{d}x = 1 - \frac{2}{3} = \frac{1}{3}$$

and

$$\mathbb{E}(X^2) = \int_0^1 x^2 \cdot 2(1 - x)\,\mathrm{d}x = \int_0^1 (2x^2 - 2x^3)\,\mathrm{d}x = \frac{2}{3} - \frac{2}{4} = \frac{1}{6}$$

so that

$$\text{var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \frac{1}{6} - \left(\frac{1}{3}\right)^2 = \frac{1}{18}.$$

Let $X_1, \ldots, X_{100}$ denote the x-coordinates of the 100 points chosen independently and uniformly at random from the triangle, and set

$$\overline{X} = \frac{X_1 + \cdots + X_{100}}{100}.$$

We are interested in computing

$$\mathbf{P}\left\{\overline{X} > \frac{7}{18}\right\}.$$

The central limit theorem tells us that

$$\frac{\overline{X} - \mathbb{E}(\overline{X})}{\sqrt{\mathrm{var}(\overline{X})}}$$

has an approximately $\mathcal{N}(0,1)$ distribution. Thus, since

$$\mathbb{E}(\overline{X}) = \mathbb{E}(X_1) = \frac{1}{3}$$

and

$$\mathrm{var}(\overline{X}) = \frac{\mathrm{var}(X_1)}{100} = \frac{1/18}{100} = \frac{1}{1800},$$

we find

$$\mathbf{P}\left\{\overline{X} > \frac{7}{18}\right\} = \mathbf{P}\left\{\frac{\overline{X} - \frac{1}{3}}{\sqrt{\frac{1}{1800}}} > \frac{\frac{7}{18} - \frac{1}{3}}{\sqrt{\frac{1}{1800}}}\right\} \approx \mathbf{P}\left\{Z > \frac{10}{\sqrt{18}}\right\}$$

where $Z \sim \mathcal{N}(0,1)$. Using a table of normal probabilities, we find

$$\mathbf{P}\left\{\overline{X} > \frac{7}{18}\right\} \approx \mathbf{P}\left\{Z > \frac{10}{\sqrt{18}}\right\} \doteq \mathbf{P}\{Z > 2.357\} \doteq 1 - 0.9909 = 0.0091.$$

6. Let X denote the number of kernels that pop so that X has a binomial distribution with parameters n and p. We now need to find n and p. We are told that $\mathbb{E}(X) = 63$ and $\mathrm{var}(X) = 6.3$. We know that, in general, if $X \sim \mathrm{Bin}(n,p)$, then $\mathbb{E}(X) = np$ and $\mathrm{var}(X) = np(1-p)$. Thus, we must solve

$$np = 63 \quad \text{and} \quad np(1-p) = 6.3$$

for n and p. Since $np = 63$, we find

$$6.3 = np(1-p) = 63(1-p) \quad \text{so that} \quad 1 - p = \frac{1}{10} \quad \text{and so} \quad p = \frac{9}{10}.$$

Thus,

$$63 = np = \frac{9n}{10} \quad \text{implies} \quad n = 700.$$

Hence, $X \sim \mathrm{Bin}(n = 700, p = 0.9)$.

7. (a) Let X denote the number of people in the sample that support Party A so that the exact distribution of X is $X \sim \mathrm{Bin}(n = 1350, p = 0.40)$. Using the normal approximation to the binomial distribution, we know that

$$\frac{X - np}{\sqrt{np(1-p)}}$$

has an approximately $\mathcal{N}(0,1)$ distribution.

Using the continuity correction, we find

$$\mathbf{P}\{X > 544\} = \mathbf{P}\{X \geq 545\} = \mathbf{P}\{X \geq 544.5\} = \mathbf{P}\left\{\frac{X - 1350 \cdot 0.4}{\sqrt{1350 \cdot 0.4 \cdot 0.6}} \geq \frac{544.5 - 1350 \cdot 0.4}{\sqrt{1350 \cdot 0.4 \cdot 0.6}}\right\}$$
$$\approx \mathbf{P}\{Z \geq 0.25\}$$

where $Z \sim \mathcal{N}(0,1)$. Using a table of normal probabilities, we find

$$\mathbf{P}\{X > 544\} \approx \mathbf{P}\{Z \geq 0.25\} \doteq 1 - 0.5987 = 0.4013.$$

7. (b) Let X denote the number of people in the sample that support Party B so that the exact distribution of X is $X \sim \text{Bin}(n = 1350, p = 0.02)$. Since $np = 1350 \cdot 0.02 = 2.7$ is quite small, the normal approximation to the binomial distribution is not applicable. Instead, we use the Poisson approximation to the binomial distribution. That is, X has an approximately Poisson($\lambda = np$) distribution. Thus,

$$\mathbf{P}\{X \leq 2\} \approx \mathbf{P}\{Y \leq 2\}$$

where $Y \sim \text{Poisson}(2.7)$, and so we find

$$\begin{aligned}
\mathbf{P}\{X \leq 2\} \approx \mathbf{P}\{Y \leq 2\} &= \mathbf{P}\{Y = 0\} + \mathbf{P}\{Y = 1\} + \mathbf{P}\{Y = 2\} \\
&= \frac{2.7^0 e^{-2.7}}{0!} + \frac{2.7^1 e^{-2.7}}{1!} + \frac{2.7^2 e^{-2.7}}{2!} \\
&= e^{-2.7}\left(1 + 2.7 + \frac{2.7^2}{2}\right) \\
&\doteq 0.4936245.
\end{aligned}$$