Math 5651: Homework 3

Matthew Pipes
February 21, 2026

Proofs for Pascal's Identity and the Binomial Theorem

Theorem: Pascal's Identity: For any integers n and r such that $1 \leq r \leq n$:

$$\binom{n}{r-1} + \binom{n}{r} = \binom{n+1}{r}$$

Proof:

$$\begin{aligned}
\binom{n}{r-1} + \binom{n}{r} &= \frac{n!}{(r-1)!(n-r+1)!} + \frac{n!}{r!(n-r)!} \\
&= \frac{r \cdot n!}{r!(n-r+1)!} + \frac{n!(n-r+1)}{r!(n-r+1)!} \\
&= \frac{r \cdot n! + n!(n-r+1)}{r!(n-r+1)!} \\
&= \frac{n!(r+n-r+1)}{r!(n-r+1)!} \\
&= \frac{n!(n+1)}{r!(n-r+1)!} \\
&= \frac{(n+1)!}{r!(n-r+1)!} \\
&= \binom{n+1}{r}
\end{aligned}$$

Proof of the Binomial Theorem by Induction

Theorem: For any positive integer n:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k$$

Proof: We proceed with induction on n.

1. Base Case

Let $n = 1$:

$$(x+y)^1 = x + y = \binom{1}{0} x^1 y^0 + \binom{1}{1} x^0 y^1 = x + y$$

The base case holds.

2. Inductive Hypothesis

Assume the theorem holds for some positive integer n:

$$(x+y)^n = \binom{n}{0} x^n + \binom{n}{1} x^{n-1} y + \cdots + \binom{n}{n-1} x y^{n-1} + \binom{n}{n} y^n \quad \text{— (1)}$$

3. Inductive Step

We must show that the theorem holds for $n+1$:

$$(x+y)^{n+1} = \sum_{k=0}^{n+1} \binom{n+1}{k} x^{n+1-k} y^k$$

First, we utilize the exponent law $(x+y)^{n+1} = (x+y)^1 (x+y)^n$ and factor the expression:

$$\begin{aligned} (x+y)^{n+1} &= (x+y)(x+y)^n \\ &= x(x+y)^n + y(x+y)^n \end{aligned}$$

Next, we substitute the inductive hypothesis (1) into this factored expression. Multiplying the expansion by x yields:

$$x(x+y)^n = \binom{n}{0} x^{n+1} + \binom{n}{1} x^n y + \cdots + \binom{n}{n-1} x^2 y^{n-1} + \binom{n}{n} x y^n \quad \text{— (2)}$$

Similarly, multiplying the expansion by y yields:

$$y(x+y)^n = \binom{n}{0} x^n y + \binom{n}{1} x^{n-1} y^2 + \cdots + \binom{n}{n-1} x y^n + \binom{n}{n} y^{n+1} \quad \text{— (3)}$$

Adding (2) and (3) together to find $(x+y)^{n+1}$, we group terms with like powers:

$$\begin{aligned}(x+y)^{n+1} = \binom{n}{0}x^{n+1} + \left[\binom{n}{0} + \binom{n}{1}\right] x^n y + \dots \\ + \left[\binom{n}{n-1} + \binom{n}{n}\right] xy^n + \binom{n}{n} y^{n+1}\end{aligned}$$

Applying **Pascal's Identity**, $\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}$, we substitute the bracketed terms:

$$(x+y)^{n+1} = \binom{n}{0}x^{n+1} + \binom{n+1}{1}x^n y + \dots + \binom{n+1}{n}xy^n + \binom{n}{n}y^{n+1}$$

Finally, since $\binom{n}{0} = \binom{n+1}{0} = 1$ and $\binom{n}{n} = \binom{n+1}{n+1} = 1$, we can write:

$$(x+y)^{n+1} = \binom{n+1}{0}x^{n+1} + \binom{n+1}{1}x^n y + \dots + \binom{n+1}{n}xy^n + \binom{n+1}{n+1}y^{n+1}$$

The theorem holds for $n+1$, and thus by the principle of mathematical induction, the Binomial Theorem holds for all positive integers n. $\blacksquare$

Problem 1

Let $X \sim \text{Poi}(\lambda)$ and $Y \sim \text{Poi}(\mu)$, show that $X + Y \sim \text{Poi}(\lambda + \mu)$

Solution

We first notice that, as X and Y Poisson random variables, they take on discrete, non negative values. If we want to find all of the ways in which $X + Y$ can equal some k, we consider that $X \in [0, i]$ and $Y \in [0, k - i]$ such that $X + Y = i + (k - i) = k$.

Now, we consider the Probability Mass Function $P(X + Y = k)$. Since X and Y are independent, we can simply multiply their individual Probability Mass Functions, as follows:

$$P(X + Y = k) = \sum_{i=0}^{k} \left(\frac{e^{-\lambda} \lambda^i}{i!} \right) \left(\frac{e^{-\mu} \mu^{(k-i)}}{(k-i)!} \right)$$

Since our summation is not dependent on $e^{-\lambda}$ and $e^{-\mu}$, we move them out of the summation:

$$= e^{-\lambda} e^{-\mu} \sum_{i=0}^{k} \frac{1}{i!(k-i)!} \cdot \lambda^i \cdot \mu^{(k-i)}$$

Notice that $e^{-\lambda} e^{-\mu}$ is equivalent to $e^{-(\lambda+\mu)}$, and that we have $\frac{1}{i!(k-i)!}$ that is nearly equivalent to the definition of a binomial coefficient. Thus we strategically multiply by $\frac{k!}{k!}$:

$$= \frac{e^{-(\lambda+\mu)}}{k!} \sum_{i=0}^{k} \frac{k!}{i!(k-i)!} \cdot \lambda^i \cdot \mu^{(k-i)}$$

$$= \frac{e^{-(\lambda+\mu)}}{k!} \sum_{i=0}^{k} \binom{k}{i} \cdot \lambda^i \cdot \mu^{(k-i)}$$

Notice now that our sum is simply the definition of the binomial theorem for $(\lambda + \mu)^k$. Thus we can substitute this in, yielding:

$$P(X + Y = k) = \frac{e^{-(\lambda+\mu)} \cdot (\lambda + \mu)^k}{k!}$$

Thus, $X + Y \sim Poi(\lambda + \mu)$, as desired.

Problem 2

Let $X \sim \text{Poi}(\lambda)$. Show that the value k which is most likely (the mode) is either rounded up or rounded down by showing:

(a) If $k < \lambda$, then $P(X = k - 1) < P(X = k)$

(b) If $k > \lambda$, then $P(X = k) > P(X = k + 1)$

Solution

Given that $X \sim \text{Poi}(\lambda)$, the random variable X can take on only discrete, non-negative integer values. We are given that $X = k$ is the point at which $P(X = k)$ has the highest probability, referred to as the mode. We are asked to prove the following two relationships between k and λ:

1. When $k < \lambda$, show that $P(X = k)$ for $k \in \{0, 1, \ldots, \lfloor \lambda \rfloor\}$ is strictly increasing.
2. When $k > \lambda$, show that $P(X = k)$ for $k \in \{\lceil \lambda \rceil, \lceil \lambda \rceil + 1, \ldots, \infty\}$ is strictly decreasing.

We must show from these relationships the mode k is either $\lfloor \lambda \rfloor$ or $\lceil \lambda \rceil$.

To find the relationship between successive probabilities, we consider the ratio:

$$\frac{P(X = k)}{P(X = k - 1)} = \frac{\frac{e^{-\lambda}\lambda^k}{k!}}{\frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!}}$$

Canceling the common term $e^{-\lambda}$ and simplifying the powers of λ and the factorials:

$$\frac{P(X = k)}{P(X = k - 1)} = \frac{\lambda^k}{\lambda^{k-1}} \cdot \frac{(k-1)!}{k!} = \lambda \cdot \frac{(k-1)!}{k(k-1)!}$$

Which simplifies to the final relationship:

$$\frac{P(X = k)}{P(X = k - 1)} = \frac{\lambda}{k}$$

From this result we notice that there is a third case that the problem doesn't explicitly ask us to solve, the case in which $\lambda = k$. As a result, the terminology "strictly increasing/decreasing" that we inferred from the problem earlier will change slightly. Let us consider the ratio $\frac{\lambda}{k}$ for our three cases:

- For all valid $k \leq \lambda$, the ratio $\frac{\lambda}{k} \geq 1$, implying $P(X = k) \geq P(X = k - 1)$. The sequence is non-decreasing.
- For all valid $k > \lambda$, the ratio $\frac{\lambda}{k} < 1$, implying $P(X = k) < P(X = k - 1)$. The sequence is strictly decreasing.
- If $k = \lambda$, the ratio $\frac{\lambda}{k} = 1$, implying $P(X = k) = P(X = k - 1)$. Thus, there are two modes at $k = \lambda$ and $k = \lambda - 1$.

Thus, in the case where $\lambda \notin \mathbb{Z}$, the probability is non-decreasing for $k \leq \lambda$ and strictly decreasing for $k > \lambda$, thus the mode occurs at $\lfloor \lambda \rfloor$.

In the case where $\lambda \in \mathbb{Z}$, the ratio $\frac{\lambda}{k}$ is exactly 1 at $\lambda = k$, resulting in two modes at λ and $\lambda - 1$.

Problem 3a

Given that $X \sim Uni(0,1)$ and $Y \sim Uni(0,1)$, we know:

$$f_X(x) = 1 \quad \text{and} \quad f_Y(y) = 1$$

Since X and Y are independent, we can find $f(x,y)$:

$$f(x,y) = f_X(x) \cdot f_Y(y) = 1 \cdot 1 = 1$$

Finally, we integrate over the relevant region:

$$\begin{aligned}
P(X > Y) &= \iint f(x,y)\,dA \\
&= \int_0^{x=1} \int_0^{y=x} 1\,dy\,dx \\
&= \int_0^{x=1} [y]_0^x\,dx \\
&= \int_0^1 x\,dx \\
&= \left.\frac{x^2}{2}\right|_0^1 \\
&= \boxed{\frac{1}{2}}
\end{aligned}$$

Problem 3b

$$\begin{aligned}
f_{X|X>Y}(x) &= \frac{f_{X,\{X>Y\}}(x)}{P(X > Y)} \\
&= \frac{f_{X,\{X>Y\}}(x)}{1/2} \\
&= 2\int_0^{y=x} 1\,dy \\
&= \boxed{2x}, \quad 0 < x < 1
\end{aligned}$$

Problem 3c

$$
\begin{aligned}
E(X \mid X > Y) &= \int_0^1 x \cdot f_{X|X>Y}(x)\,dx \\
&= \int_0^1 x \cdot (2x)\,dx \\
&= \int_0^1 2x^2\,dx \\
&= \left[\frac{2}{3}x^3\right]_0^1 \\
&= \boxed{\frac{2}{3}}
\end{aligned}
$$

Problem 4a

$$f(x,y) = \frac{6}{7}(x^2 + \frac{xy}{2}) \quad 0 < x < 1; \quad 0 < y < 2$$

$$\frac{6}{7}\int_0^{x=1}\int_0^{y=2} x^2 + \frac{xy}{2}dydx = \frac{6}{7}\int_0^{x=1} yx^2 + \frac{xy^2}{4}\bigg|_0^2 dx = \frac{6}{7}\int_0^{x=1} 2x^2 + xdx$$

$$= \frac{6}{7}\left(\frac{2x^3}{3} + \frac{x^2}{2}\right)\bigg|_0^1 = \frac{6}{7}\left(\frac{2}{3} + \frac{1}{2}\right) = \frac{6}{7}\left(\frac{4}{6} + \frac{3}{6}\right)$$

$$= \frac{6}{7}\left(\frac{7}{6}\right) = 1$$

Problem 4b

$$f_x(x) = \frac{6}{7}\int_0^{y=2}(x^2 + \frac{xy}{2})dy = \frac{6}{7}(x^2y + \frac{xy^2}{4})\bigg|_0^2 = \frac{6}{7}(2x^2 + x)$$

Problem 4c

$$f_y(y) = \frac{6}{7}\int_0^{x=1}(x^2 + \frac{xy}{2})dx = \frac{6}{7}(\frac{x^3}{3} + \frac{x^2y}{4})\bigg|_0^{x=1} = \frac{6}{7}(\frac{1}{3} + \frac{y}{4})$$

Problem 4d

$$E[x] = \frac{6}{7}\int_0^{x=1} x(2x^2 + x)dx = \frac{6}{7}\int_0^{x=1} 2x^3 + x^2dx = \frac{6}{7}(\frac{x^4}{2} + \frac{x^3}{3})\bigg|_0^{x=1}$$

$$= \frac{6}{7}(\frac{1}{2} + \frac{1}{3}) = \frac{6}{7}(\frac{3}{6} + \frac{2}{6}) = \frac{6}{7}(\frac{5}{6}) = \frac{5}{7}$$

Problem 4e

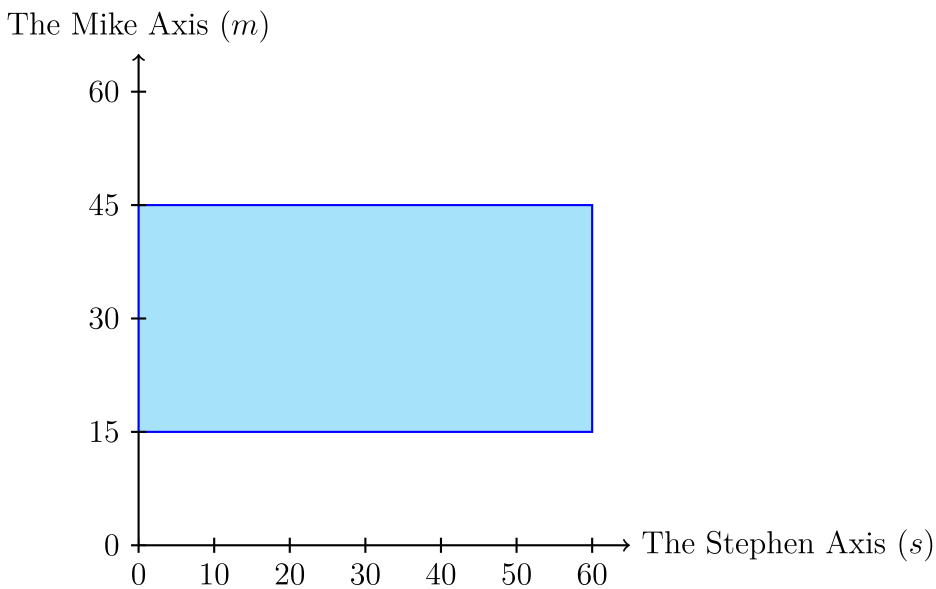
$$E(y) = \frac{6}{7}\int_0^{y=2} y(\frac{1}{3} + \frac{y}{4})dy = \frac{6}{7}\int_0^{y=2} \frac{y}{3} + \frac{y^2}{4}dy = \frac{6}{7}(\frac{y^2}{6} + \frac{y^3}{12})\bigg|_0^{y=2}$$

$$= \frac{6}{7}(\frac{4}{6} + \frac{8}{12}) = \frac{6}{7}(\frac{2}{3} + \frac{2}{3}) = \frac{6}{7}(\frac{4}{3}) = \frac{8}{7}$$

Problem 5

5. Mike and Stephen agree to meet in Central Park after lunch. Let the number of minutes after 12:00 noon that Stephen arrives be $S \sim \text{Uni}(0, 60)$ and let the number of minutes after 12:00 noon that Mike arrives be $M \sim \text{Uni}(15, 45)$. In other words, Stephen arrives (uniform) randomly between 12:00 and 1:00 and Mike arrives (uniform) randomly between 12:15 and 12:45. What is the probability that whoever arrives first has to wait no longer than 5 minutes?

Solution



We begin by noting a few obvious facts about this joint probability distribution. Since $M \sim \text{Uni}[15, 45]$ and $S \sim \text{Uni}[0, 60]$, we know that their marginal probability distributions are $f_S(s) = \frac{1}{60}$ and $f_M(m) = \frac{1}{30}$. As S and M are independent, we calculate their joint probability distribution to be $f(s, m) = \frac{1}{1800}$. This is verified easily by the figure showing the sample space for our distributions above.

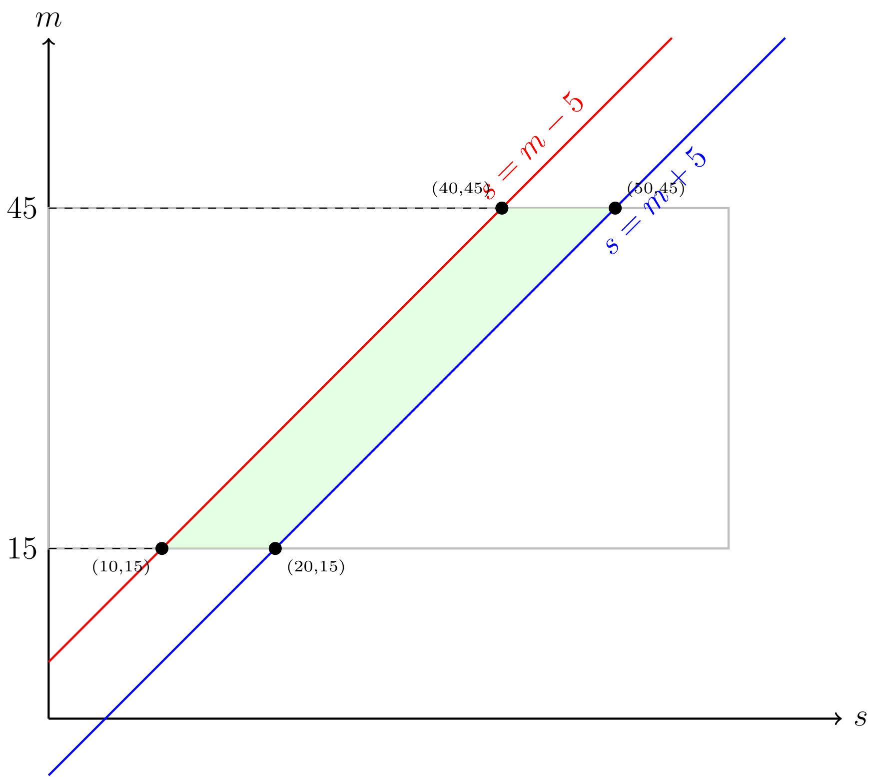
Ignoring probabilities for a bit, and focusing on the geometry of the problem, we must find an area for which $|S - M \leq 5|$. Expanding the algebra and solving for S yields the two following inequalities:

$$\begin{cases} S \leq m + 5 \\ S \geq m - 5 \end{cases}$$

To find the vertices of the success region within our sample space, we evaluate our system of inequalities at the boundaries of Mike's arrival interval, $M \in [15, 45]$:

$$\begin{aligned} \text{For } M = 15: \quad & S = 15 - 5 = 10 \implies (10, 15) \\ & S = 15 + 5 = 20 \implies (20, 15) \\[1ex] \text{For } M = 45: \quad & S = 45 - 5 = 40 \implies (40, 45) \\ & S = 45 + 5 = 50 \implies (50, 45) \end{aligned}$$

Combining these boundary points with the inequalities we derived earlier yields the following:



Finally, we have all the tools we need to solve the problem using probability. We have reduced the problem to the following double integral:

$$P(|S - M| \leq 5) = \int_{15}^{45} \left(\int_{m-5}^{m+5} \frac{1}{1800} \, ds \right) dm$$

Note our choice of integration over the Stephen axis first requires only one integral, whereas integration over the Mike axis would require three sub-integrals.

All that's left to do now is to evaluate:

$$\begin{aligned}
P(|S-M| \le 5) &= \int_{15}^{45} \left[\frac{s}{1800}\right]_{m-5}^{m+5} dm \\
&= \int_{15}^{45} \left(\frac{(m+5)-(m-5)}{1800}\right) dm \\
&= \int_{15}^{45} \frac{10}{1800}\, dm \\
&= \int_{15}^{45} \frac{1}{180}\, dm \\
&= \left[\frac{m}{180}\right]_{15}^{45} \\
&= \frac{45}{180} - \frac{15}{180} \\
&= \frac{30}{180} \\
&= \frac{1}{6}
\end{aligned}$$

Thus, there is a $\frac{1}{6}$ probability that Stephen and Mike will arrive within 5 minutes of each other.