MATH 340 – SPRING 2026 – HOMEWORK 3

Due Thursday, January 29, 2026 at 8am on Gradescope. You must justify all of your answers for full credit.

You are encouraged to collaborate with other students, but you must write up your solutions individually, without reference to notes from the collaboration. You may not search the internet or ask AI for solutions to the homework problems. Exception: it is fine to use AI to "vibe-code" the programming questions if you want, but make sure you understand what the code is doing and how to modify it.

Problem 1.

(a) Suppose that $\Omega = \Omega_1 \times \cdots \times \Omega_n$ and $p(\omega_1, \ldots, \omega_n) = p_1(\omega_1) \cdots p_n(\omega_n)$. Let $X_i(\omega) = f_i(\omega_i)$ for some function f_i. Prove that $X_1, \ldots, X_n$ are independent random variables.

(b) Let X_1, X_2, X_3 be independent random variables, and let $g\colon \mathbb{R}^2 \to \mathbb{R}$ and $h\colon \mathbb{R} \to \mathbb{R}$ be functions. Show that $g(X_1, X_2)$ and $h(X_3)$ are independent. State and prove a generalization of this result.

Solution.

(a) Let $x_1, \ldots, x_n \in \mathbb{R}$. Then we have

$$\begin{aligned}
P(X_1 \le x_1, \ldots, X_n \le x_n) &= P\left(\{(\omega_1, \ldots, \omega_n) \in \Omega : f_i(\omega_i) \le x_i \text{ for each } i\}\right) \\
&= P\left(\bigtimes_{i=1}^{n} \{\omega_i \in \Omega_i : f_i(\omega_i) \le x_i\}\right) \\
&= \sum_{\omega_1 \in \Omega_n} \cdots \sum_{\omega_n \in \Omega_n} \prod_{i=1}^{n} \left(p_i(\omega_i)\mathbf{1}\{f_i(\omega_i) \le x_i\}\right) \\
&= \prod_{i=1}^{n} \sum_{\omega_i \in \Omega_n} p_i(\omega_i)\mathbf{1}\{f_i(\omega_i) \le x_i\} \\
&= \prod_{i=1}^{n} P(X_i \le x_i),
\end{aligned}$$

which means that $X_1, \ldots, X_n$ are independent.

(b) A generalization of this could be that if $X_{1,1}, \ldots, X_{1,n_1}, X_{2,1}, \ldots, X_{2,n_2}, \ldots, X_{m,1}, \ldots, X_{m,n_m}$ are independent random variables, and $g_i\colon \mathbb{R}^{n_i} \to \mathbb{R}$ are functions, then

$$g_1(X_{1,1}, \ldots, X_{1,n_1}), \ldots, g_m(X_{m,1}, \ldots, X_{m,n_m})$$

are also independent random variables. To prove this, we write

$$P(g_1(X_{1,1},\dots,X_{1,n_1}) = x_1,\dots,g_m(X_{m,1},\dots,X_{m,n_m}) = x_m)$$
$$= \sum_{x_{1,1},\dots,x_{1,n_1}} \cdots \sum_{x_{m,1},\dots,x_{m,n_m}} P(X_{i,j} = x_{ij} \text{ for all } i,j) \prod_{i=1}^{m} \mathbf{1}\{g_i(x_{i,1},\dots,x_{i,n_i}) = x_i\}$$
$$= \sum_{x_{1,1},\dots,x_{1,n_1}} \cdots \sum_{x_{m,1},\dots,x_{m,n_m}} \left(\prod_{i,j} P(X_{i,j} = x_{i,j})\right) \prod_{i=1}^{m} \mathbf{1}\{g_i(x_{i,1},\dots,x_{i,n_i}) = x_i\}$$
$$= \sum_{x_{1,1},\dots,x_{1,n_1}} \cdots \sum_{x_{m,1},\dots,x_{m,n_m}} \prod_{i=1}^{m} \left(\prod_{j=1}^{n_i} P(X_{i,j} = x_{i,j})\right) \mathbf{1}\{g_i(x_{i,1},\dots,x_{i,n_i}) = x_i\}$$
$$= \sum_{x_{1,1},\dots,x_{1,n_1}} \cdots \sum_{x_{m,1},\dots,x_{m,n_m}} \prod_{i=1}^{m} \left(P\left(X_{i,j} = x_{i,j} \text{ for all } j = 1,\dots,n_i\right) \mathbf{1}\{g_i(x_{i,1},\dots,x_{i,n_i}) = x_i\}\right)$$
$$= \prod_{i=1}^{m} P(g_i(X_{i,1},\dots,X_{i,n_i}) = x_i),$$

which is what we needed to show.

Problem 2. Let $\Omega = \{0,1,2,3\}$ be a sample space equipped with probability measure given by $p(0) = 1/2$, $p(1) = 1/6$, $p(2) = 1/6$, $p(3) = 1/6$.

(a) Exhibit two random variables X and Y defined with respect to the sample space Ω such that $X \sim \text{Ber}(1/3)$ and $Y \sim \text{Ber}(1/3)$ but X and Y are not the same random variable.
(b) Show that there is no random variable X defined on Ω such that $X \sim \text{Bin}(2,1/2)$.
(c) Show that if three random variables X,Y,Z on Ω are independent, then there must be some $x \in \mathbb{R}$ such that either $P(X = x) = 1$, $P(Y = x) = 1$, or $P(Z = x) = 1$.

Solution.

(a) We could take $X(0) = X(1) = 0$ and $X(2) = X(3) = 1$, and $Y(0) = Y(2) = 0$ and $Y(3) = Y(4) = 1$. Then it's easy to check the desired conditions.
(b) To have $X \sim \text{Bin}(2,1/2)$, we must have $P(X = 0) = 1/4$. But no subset of Ω has probably $1/4$, as all possible probabilities are multiples of $1/6$.
(c) If not, then there must be $x_1,x_2,y_1,y_2,z_1,z_2 \in \mathbb{R}$ such that $P(X = x_1), P(X = x_2), P(Y = y_1), P(Y = y_2), P(Z = z_1), P(Z = z_2) > 0$. By independence, this must mean that for any $(a,b,c) \in \{x_1,x_2\} \times \{y_1,y_2\} \times \{z_1,z_2\}$, we have $P(X = a, Y = b, Z = c) > 0$. This means that there must be at least eight possible outcomes of the experiment because there are at least eight possible values of the random variables, but Ω only has four elements.

Problem 3. Recall the airplane problem from Homework 1, Problem 3. Let X_n be the number of passengers who sit in their assigned seat. For each $k \geq 0$, compute $\lim_{n\to\infty} p_{X_n}(k)$. [*Hint*: if exactly k people sit in their assigned seats, then none of the remaining $n-k$ people can be sitting in their assigned seats, so you can use the result of Homework 1, Problem 3 with n replaced by $n-k$.]

Solution. Let σ_i be the seat of the ith passenger. We have

$$P(X_n = k) = \sum_{\{\ell_1,\dots,\ell_k\} \in \binom{\{1,\dots,n\}}{k}} P\left(\sigma_j = j \text{ for } j \in \{\ell_1,\dots,\ell_k\} \text{ and } \sigma_j \neq j \text{ for } j \notin \{\ell_1,\dots,\ell_k\}\right)$$
$$= \sum_{\{\ell_1,\dots,\ell_k\} \in \binom{\{1,\dots,n\}}{k}} \frac{(n-k)!}{n!} \cdot P\left(\sigma_j \neq j \text{ for } j \notin \{\ell_1,\dots,\ell_k\} \mid \sigma_j = j \text{ for } j \in \{\ell_1,\dots,\ell_k\}\right).$$

The last probability is equal to the probability that no one on a plane with $n-k$ passengers sits in their assigned seat, which we computed before is $1 - \sum_{\ell=1}^{n-k} \frac{(-1)^{\ell+1}}{\ell!} = \sum_{\ell=0}^{n-k} \frac{(-1)^\ell}{\ell!}$. So we get

$$P(X_n = k) = \binom{n}{k} \frac{(n-k)!}{n!} \sum_{\ell=0}^{n-k} \frac{(-1)^\ell}{\ell!} = \frac{1}{k!} \sum_{\ell=0}^{n-k} \frac{(-1)^\ell}{\ell!}.$$

As $n \to \infty$, this quantity approaches $\mathrm{e}^{-1}/k!$, which is the pmf of a Poisson(1) random variable.

Problem 4. Suppose that $X \sim \text{Bin}(n, p)$ and $Y \sim \text{Bin}(m, p)$ are independent. Compute (with proof) the pmf of the random variable $X + Y$.

Proof. We can realize X and Y on a single sample space Ω by letting $\Omega = \{0,1\}^{m+n}$ with probability measure $p(\omega) = p^{\omega_1 + \cdots + \omega_{n+m}}(1-p)^{n-(\omega_1+\cdots+\omega_{n+m})}$, and then setting $X(\omega) = \omega_1 + \cdots + \omega_n$ and $Y(\omega) = \omega_{n+1} + \cdots + \omega_{m+n}$. Then $(X+Y)(\omega) = \omega_1 + \cdots + \omega_{m+n}$ and so $X + Y \sim \text{Bin}(m+n, p)$. Hence we have $p_{X+Y}(k) = \binom{m+n}{k} p^k (1-p)^{m+n-k}$ for $k = 0, \ldots, m+n$. □

Problem 5 (some problems about the geometric distribution)**.**

(a) Let $X \sim \text{Geom}(p)$. Prove that $P(X = m + k \mid X > m) = P(X = k)$. Explain why this is called the *memoryless property* of the geometric distribution.

(b) Let $X_1, \ldots, X_r \sim \text{Geom}(p)$ be independent. Show that $X_1 + \cdots + X_r$ has a negative binomial distribution with parameters (p, r).

Proof.

(a) We have that

$$P(X > m) = \sum_{k=m+1}^{\infty} p(1-p)^{k-1} = p \cdot \frac{(1-p)^m}{1-(1-p)} = (1-p)^m,$$

and hence that

$$P(X = m + k \mid X > m) = \frac{P(X = m+k)}{P(X > m)} = \frac{p(1-p)^{m+k-1}}{(1-p)^m} = p(1-p)^{k-1} = P(X = k).$$

This is called the memoryless property because, given that a failure hasn't happened yet by time m, the distribution of the additional time you will have to wait is the same as the distribution of the waiting time from the start. So the system doesn't "remember" that you have already waited for time m.

(b) We argue by induction on r. For $r = 1$, we simply observe that the formulas for the pmfs match. So suppose the statement is true for r, and we'll try to prove it for $r + 1$. Let X be a Geom(p) random variable, let Y be an independent negative binomial random variable with parameters (p, r), and let $Z = X + Y$. Then we have

$$\begin{aligned} P(Z = z) &= \sum_x P(X = x, Y = z - x) \\ &= \sum_x P(X = x) P(Y = z - x) \\ &= \sum_{x=1}^{z-r} p(1-p)^{x-1} \binom{z-x-1}{r-1} p^r (1-p)^{z-x-r} \\ &= p^{r+1}(1-p)^{z-r-1} \sum_{x=1}^{z-r} \binom{z-x-1}{r-1}. \end{aligned}$$

To compute the last sum, we count the number of ways of picking r elements out of $z-1$ elements in two different ways. The first is to simply write it as $\binom{z-1}{r}$. The second is to condition on the first element that is chosen, which we call x. The first element can be any of the first $z-r$ elements, and once we have chosen that element to be x, there are $\binom{z-x-1}{r-1}$ choices of the remaining $r-1$ elements. So we get $\binom{z-1}{r} = \sum_{x=1}^{z-r}\binom{z-x-1}{r-1}$. Using this in the above, we get $P(Z=z) = \binom{z-1}{r}p^{r+1}(1-p)^{z-r-1}$, which is what we needed to show.

□

Problem 6. Let $s>1$. Let X be a random variable such that $p_X(k) = k^{-s}/\zeta(s)$ for $k=1,2,3,\ldots$ and $p_X(x)=0$ for other $x\in\mathbb{R}$. Here the normalizing factor is

$$\zeta(s) \coloneqq \sum_{k=1}^{\infty} k^{-s}.$$

This function is also known as the *Riemann zeta function* and it is very important in number theory. For each prime number q, let Y_q be the random variable such that the prime factorization of X (which is unique) can be written as

$$X = \prod_{q \text{ prime}} q^{Y_q}.$$

(a) Show that $Y_q + 1 \sim \mathrm{Geom}(p_q)$ and compute p_q.
(b) Show that the family $\{Y_q\}_{q \text{ prime}}$ is an independent family of random variables. [*Hint*: it suffices to show that for any distinct prime numbers $q_1,\ldots,q_n$ and any nonnegative integers $m_1,\ldots,m_n$, the events $\{Y_{q_1}\ge m_1\},\ldots,\{Y_{q_n}\ge m_n\}$ are independent.]
(c) By using the last two parts, along with a limit lemma proved in class, show that

$$\frac{1}{\zeta(s)} = P(X=1) = \prod_{q \text{ prime}} (1-q^{-s}).$$

This is the famous *Euler product formula* for the Riemann zeta function.

Solution.

(a) We can write, for a natural number $k\ge 1$,

$$P(Y_q+1\ge k) = P(Y_q \ge k-1) = \frac{1}{\zeta(s)}\sum_{\ell=1}^{\infty}(\ell q^{k-1})^{-s} = (q^{-s})^{k-1},$$

so, for natural numbers $k\ge 1$, we have

$$\begin{aligned} P(Y_q+1=k) &= P(Y_q+1\ge k) - P(Y_q+1\ge k+1) \\ &= (q^{-s})^{k-1} - (q^{-s})^k = (q^{-s})^{k-1}(1-q^{-s}). \end{aligned}$$

This is the pmf of a geometric distribution with parameter $p_q = 1-q^{-s}$.

(b) We have $Y_{q_1}\ge m_1,\ldots,Y_{q_n}\ge m_n$ if and only if X is divisible by $q_1^{m_1}\cdots q_n^{m_n}$. Therefore, we have

$$\begin{aligned} P(Y_{q_1}\ge m_1,\ldots,Y_{q_n}\ge m_n) &= \frac{1}{\zeta(s)}\sum_{\ell=1}^{\infty}(\ell q_1^{m_1}\cdots q_n^{m_n})^{-s} = \frac{(q_1^{m_1}\cdots q_n^{m_n})^{-s}}{\zeta(s)}\sum_{\ell=1}^{\infty}\ell^{-s} \\ &= (q_1^{m_1}\cdots q_n^{m_n})^{-s} = P(Y_{q_1}\ge m_1)\cdots P(Y_{q_n}\ge m_n), \end{aligned}$$

so the Y_{q_i}s are independent.

(c) That $\frac{1}{\zeta(s)} = P(X=1)$ is clear from the pmf of X. On the other hand, we can write

$$\{X=1\} = \bigcap_{q \text{ prime}}\{Y_q=0\} = \bigcap_{n=1}^{\infty}\left(\bigcap_{q \text{ prime}, q\le n}\{Y_q=0\}\right).$$

The sequence $\left(\bigcap_{q \text{ prime}, r \leq n} \{Y_q = 0\}\right)_n$ is a decreasing sequence of events, so by the lemma proved in class, we have

$$\begin{aligned}
P(X=1) &= \lim_{n\to\infty} P\left(\bigcap_{q \text{ prime}, q\leq n} \{Y_q = 0\}\right) \\
&= \lim_{n\to\infty} \prod_{q \text{ prime}, q\leq n} P(Y_q = 0) \\
&= \lim_{n\to\infty} \prod_{q \text{ prime}, q\leq n} (1 - q^{-s}) \\
&= \prod_{q \text{ prime}} (1 - q^{-s}),
\end{aligned}$$

as claimed.