Section 5 Discrete Random Variables

Statistics I · Fall 2026 · Zong-Wei Yeh · Department of Finance, CCU

Overview

Section 4 built the general machinery of a random variable: the distribution, the expectation and variance, and above all the moment generating function. But machinery is not a model. To say anything about a real count we need a pmf with one or two parameters that fits a recurring situation. This section gives five such models. They recur throughout statistics and finance.

This section has six parts.

- §5.1 covers the Bernoulli, a single yes/no trial. Most of the other models are built from it.
- §5.2 adds n independent trials and counts the successes. That is the binomial.
- §5.3 fixes the number of successes at one instead and counts the trials. That is the geometric, and it is memoryless.
- §5.4 drops the independence, by sampling without replacement from a finite population. That is the hypergeometric.
- §5.5 covers the Poisson, which counts events arriving at a steady rate. It is also the limit of the binomial when trials are many and successes are rare.
- §5.6 collects the relationships between the five in one map.

Two tools from Section 4 do most of the work. Where a distribution has a workable MGF, we read its mean and variance off the derivatives at 0 (Proposition 4.40). So we do not need to sum $x\,p(x)$ over every value X can take. And where a distribution is a sum of simpler independent ones, its MGF is a product (Propositions 4.63 and 4.80). That is how the binomial and the sum of Poissons are identified. One distribution has no elementary MGF, the hypergeometric. For its mean we write the count as a sum of $0/1$ variables. Section 6 does the same work for continuous variables, where the normal and the exponential play these roles.

Distribution	pmf and support	Mean	Variance	MGF
Bernoulli(p)				
Bin(n, p)				
Geometric(p)				
Hyper-geometric				
Poisson(λ)				

Fill this table in as the section goes. Every entry is derived below, except the hypergeometric variance, which §5.4 states without proof. Here $q = 1 - p$ and $\pi = K/N$.

5.1 The Bernoulli Distribution

Definition 5.1 (Bernoulli distribution).
A random variable X is Bernoulli (伯努利分配) *with parameter $p \in [0, 1]$, written $X \sim$ Bernoulli(p), if it takes only the values* 1 *("success") and* 0 *("failure") with*

$$P(X = 1) = p, \qquad P(X = 0) = q := 1 - p,$$

that is, pmf $p(x) =$ for $x \in \{0, 1\}$.

A Bernoulli variable models a single yes/no trial (試驗). Examples are a loan that defaults or not, a part that is defective or not, and an option that finishes in the money or not. The binomial, the geometric and, in the limit, the Poisson are all built from it. So its two summaries are worth having in closed form.

Proposition 5.2 (Bernoulli MGF, mean, variance).
If $X \sim$ *Bernoulli*(p) *then* $M_X(t) =$, *and hence* $\mathbb{E}(X) = p$ *and* $\mathrm{Var}(X) =$.

Proof.

□

Example 5.3. A loan defaults with probability $p = 0.02$; let $X = 1$ if it defaults. Write the pmf and find $\mathbb{E}(X)$, $\mathrm{Var}(X)$, and σ_X.

5.2 The Binomial Distribution

Definition 5.4 (Binomial distribution).
Let X *count the successes in* n independent *Bernoulli*(p) *trials. Then* X *is* binomial (二項分配), *written* $X \sim Bin(n, p)$, *with support* (支撐集) *(the set of values* X *can take)* $\{0, 1, \ldots, n\}$ *and pmf*

$$p(x) = \qquad , \qquad x = 0, 1, \ldots, n.$$

A binomial variable adds up n independent yes/no trials. The pmf multiplies two things. The first is the probability of one particular sequence with x successes. The second counts how many such sequences there are. Figure 1 shows how the shape depends on p.

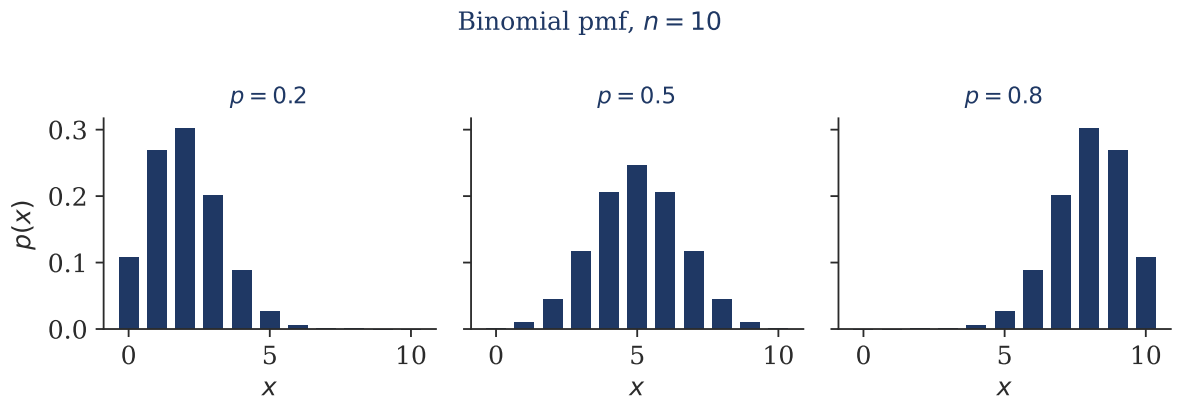


Figure 1: Binomial pmf for $n = 10$. The distribution centers near np. It is right-skewed for small p ($p = 0.2$), symmetric at $p = 0.5$, and left-skewed for large p ($p = 0.8$). The three are mirror images, because success and failure swap roles under $p \leftrightarrow q$.

Proposition 5.5 (Valid pmf)**.**
$\sum_{x=0}^{n} \binom{n}{x} p^x q^{n-x} = 1.$

Proof.

□

Proposition 5.6 (Binomial MGF)**.**
$M_X(t) =$.

Proof.

□

Proposition 5.7 (A sum of independent Bernoullis is binomial)**.**
If $X_1, \ldots, X_n$ are independent Bernoulli(p) and $S = X_1 + \cdots + X_n$, then $S \sim Bin(n, p)$.

Proof.

□

This is where the pmf in the binomial definition comes from. The counting argument under that definition and this proof are two routes to one pmf. The counting route explains the $\binom{n}{x}$. The MGF route explains why the binomial is the natural model for a total.

Proposition 5.8 (Binomial mean and variance).
$\mathbb{E}(X) =$ *and* $\mathrm{Var}(X) =$.

Proof.

□

Example 5.9. $X \sim \mathrm{Bin}(10, 0.3)$. Find $P(X = 3)$, $P(X \leq 2)$, and $\mathbb{E}(X)$, $\mathrm{Var}(X)$, σ_X.

5.3 The Geometric Distribution

Definition 5.10 (Geometric distribution).
Run independent Bernoulli(p) trials until the first success. Let X be the number of trials up to and including that success. Then X is geometric (幾何分配) *with parameter $p \in (0, 1)$, written $X \sim$ Geometric(p). It has support $\{1, 2, 3, \dots\}$ and pmf*

$$p(x) = \quad , \qquad x = 1, 2, 3, \dots$$

A geometric variable is the waiting time, in trials, for the first success. The pmf multiplies

two things. The first is the probability that the early trials all fail. The second is the probability that the last trial succeeds. Figure 2 shows the shape.

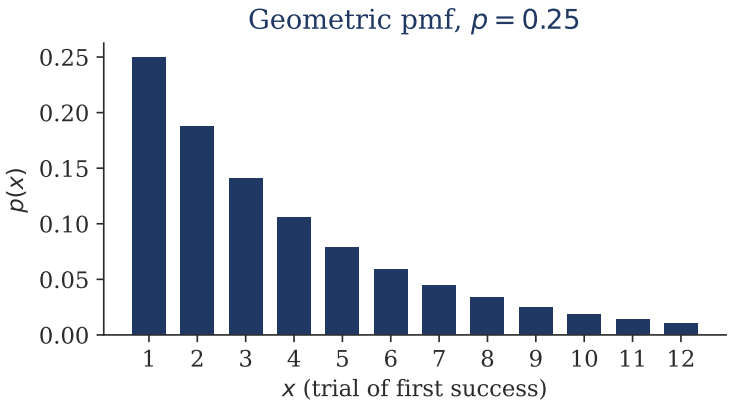


Figure 2: Geometric pmf for $p = 0.25$. Each bar is a factor $q = 0.75$ shorter than the one before it, so the probability of a long wait falls off geometrically. The most likely single outcome is always $x = 1$.

Proposition 5.11 (Valid pmf and tail)**.**
$\sum_{x=1}^{\infty} q^{x-1}p = 1$, *and* $P(X > x) =$ *for every integer* $x \geq 0$.

Proof.

□

Proposition 5.12 (Geometric MGF, mean, variance)**.**
$M_X(t) =$ *for* $qe^t < 1$. *Hence* $\mathbb{E}(X) =$ *and* $\mathrm{Var}(X) =$.

Proof.

□

Theorem 5.13 (Memorylessness)**.**
For integers $m, n \geq 0$, $P(X > m + n \mid X > m) =$.

Proof.

□

A geometric variable is *memoryless* (無記憶性). Suppose we have already waited m trials without a success. That tells us nothing about how many more we must wait. The process does not "remember" past failures, and a success is never "due". Section 6 introduces the continuous version of this property, in the exponential distribution.

Remark (alternative convention). Some texts count the number of *failures before* the first success instead. Write $Y = X - 1$. Then Y has support $\{0, 1, 2, \dots\}$, pmf $q^x p$, and mean $\mathbb{E}(Y) = q/p$. Its variance is the same, q/p^2. Both are called "geometric". This course uses the trials convention throughout, so X with mean $1/p$.

Example 5.14. Each independent drilling attempt strikes oil with probability $p = 0.25$. Let X be the attempt on which oil is first struck. Find $P(X = 3)$, $\mathbb{E}(X)$, and $P(X > 4)$.

5.4 The Hypergeometric Distribution

Definition 5.15 (Hypergeometric distribution)**.**
A population of $N \geq 2$ items contains K successes and $N - K$ failures, with $0 \leq K \leq N$. Draw n items, $1 \leq n \leq N$, without replacement*, and let X be the number of successes drawn. Then X is* hypergeometric (超幾何分配)*, written $X \sim$ Hypergeometric(N, K, n), with pmf*

$$p(x) = \qquad\qquad ,$$

for integer x with $\max(0,\ n - (N - K)) \leq x \leq \min(n, K)$.

The hypergeometric distribution is the binomial's counterpart for a finite population. Drawing without replacement makes the draws _____. Each success removed lowers the chance of the next one. So the count is no longer binomial. The pmf is a ratio of counts. The top counts the samples that contain exactly x successes. The bottom counts all samples of size n, each equally likely.

These probabilities add up to 1. That fact is a counting identity, called Vandermonde's identity. We take it as given and do not prove it here, because the course does not use the identity again. See, for example, Ross, *A First Course in Probability*.

Proposition 5.16 (Hypergeometric mean and variance)**.**
With $\pi = K/N$*,*

$$\mathbb{E}(X) = \qquad , \qquad \mathrm{Var}(X) = \qquad .$$

Proof.

□

The variance is stated without proof. Its proof uses the same indicators. It adds up their variances and the covariances between every pair of draws. Those covariances are negative. The proof is long. The later sections need only the formula. So we leave it out. See, for example, Ross, *A First Course in Probability*.

The variance ends in an extra factor that the binomial does not have. That factor is the _____. It is below 1 whenever $n \geq 2$, because dependence between draws reduces spread. It equals 1 when $n = 1$, since one draw is just a Bernoulli. It shrinks to 0 as $n \to N$, since drawing the whole population leaves nothing random.

Example 5.17. A batch of $N = 50$ parts contains $K = 20$ defectives. Inspect $n = 5$ parts without replacement; let X be the number of defectives found. Find $P(X = 2)$, $\mathbb{E}(X)$, and $\mathrm{Var}(X)$.

Remark (binomial approximation). Suppose the population is large relative to the sample. Then removing a few items barely changes its composition. So each draw behaves almost like an independent trial. The hypergeometric is then close to $\mathrm{Bin}(n, \pi)$. Precisely: let $N \to \infty$ and $K \to \infty$ with $K/N \to \pi$, while n and x stay fixed. Then $p(x) \to \binom{n}{x}\pi^x(1-\pi)^{n-x}$. The correction factor in the variance tends to 1. We state this limit without proof (see, for example, Ross, *A First Course in Probability*). The approximation improves quickly as N grows. With $n = 5$ and $\pi = 0.2$, the hypergeometric $P(X = 1)$ is 0.4313 at $N = 50$. It is 0.4148 at $N = 200$ and 0.4097 at $N = 10\,000$. The binomial gives 0.4096.

5.5 The Poisson Distribution

Many counts arise from events that happen at a steady average rate (發生率) over time or space. Examples are calls to a help desk per hour, flaws per meter of cable, and claims per month. The model behind all three is the *Poisson process* (卜瓦松過程): events arrive independently, at a steady rate λ per unit, one at a time. Under its assumptions the count over one unit follows the Poisson distribution. We state this link without proof. Proving it needs the process itself. This course does not develop the process. Theorem 5.21 below shows the idea: many small pieces of time, each with a rare event.

Definition 5.18 (Poisson distribution)**.**
X is Poisson (卜瓦松分配) *with parameter $\lambda > 0$, written $X \sim$ Poisson(λ), if it has support $\{0, 1, 2, \dots\}$ and pmf*

$$p(x) = \qquad , \qquad x = 0, 1, 2, \dots$$

The Poisson distribution has a single parameter λ. As λ grows, the distribution spreads out and becomes more symmetric (Figure 3).

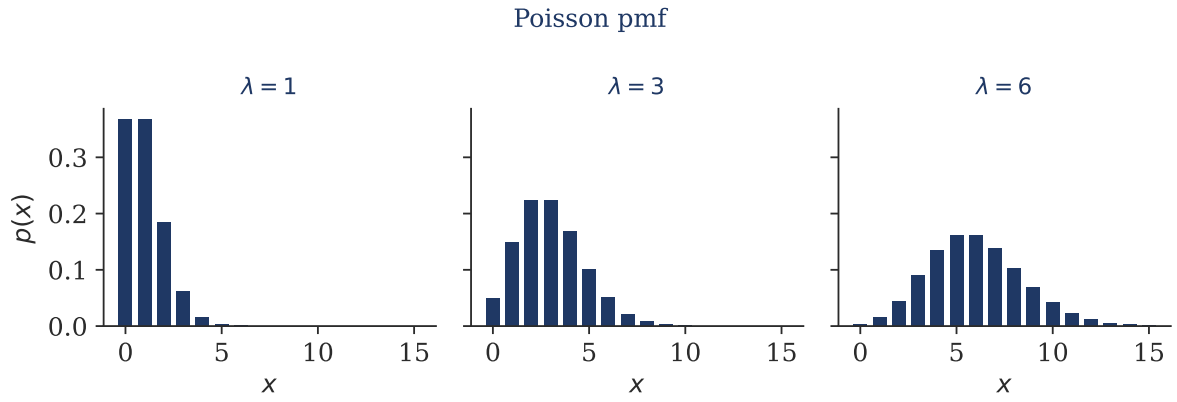


Figure 3: Poisson pmf for $\lambda = 1, 3, 6$. The peak sits near λ, and the spread grows with λ. For small λ the mass piles up near 0. For larger λ it becomes bell-shaped.

Proposition 5.19 (Valid pmf)**.**
$\sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} = 1.$

Proof.

□

Proposition 5.20 (Poisson MGF, mean, variance)**.**
$M_X(t) =$ *, and* $\mathbb{E}(X) = \mathrm{Var}(X) = \lambda.$

Proof.

□

Theorem 5.21 (Poisson limit of the binomial)**.**
Fix $\lambda > 0$ and let $n \to \infty$ with $p =$. Then for each fixed x,

$$\binom{n}{x} p^x (1-p)^{n-x} \longrightarrow \frac{e^{-\lambda}\lambda^x}{x!}.$$

Proof.

□

Proposition 5.22 (Sum of independent Poissons)**.**
If $X \sim Poisson(\lambda_1)$ and $Y \sim Poisson(\lambda_2)$ are independent, then $X + Y \sim Poisson($).

Proof.

□

Example 5.23. A help desk receives calls as a Poisson process at $\lambda = 3$ per hour. In a given hour find $P(X = 0)$, $P(X = 2)$, and $P(X \geq 1)$.

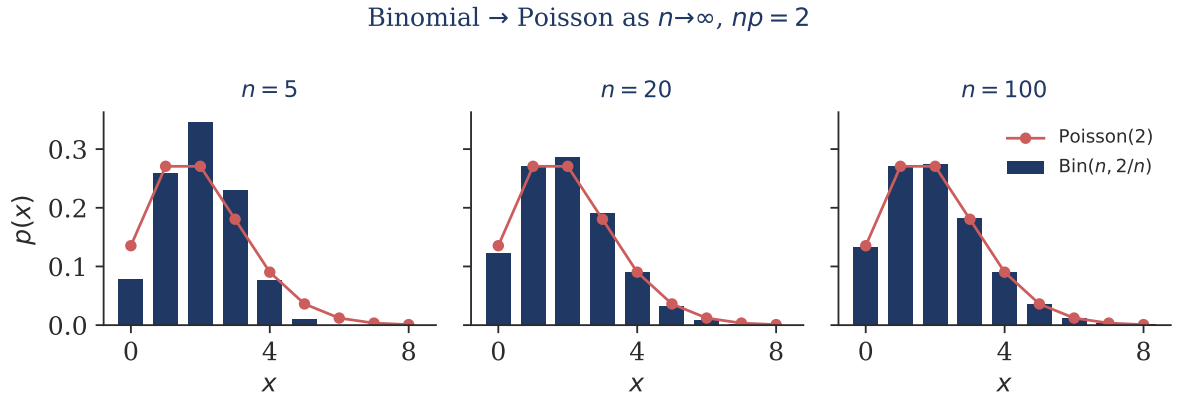


Figure 4: The binomial $\text{Bin}(n, 2/n)$ (bars) closing on the Poisson(2) pmf (red) as n grows from 5 to 100, with $np = 2$ held fixed. This is a picture of Theorem 5.21.

Example 5.24. A portfolio holds $n = 200$ policies, each lapsing independently with probability $p = 0.01$ in a year. Approximate $P(\text{no lapses})$ with a Poisson, and compare with the exact binomial.

5.6 How the Five Fit Together

Most distributions in this section start from trials with two outcomes. The differences are in what we count and how we stop. Figure 5 puts the five on one page. Every arrow on it is a definition or a result proved above, with one exception. The arrow from the hypergeometric back to the binomial is the Remark in §5.4, which we stated without proof.

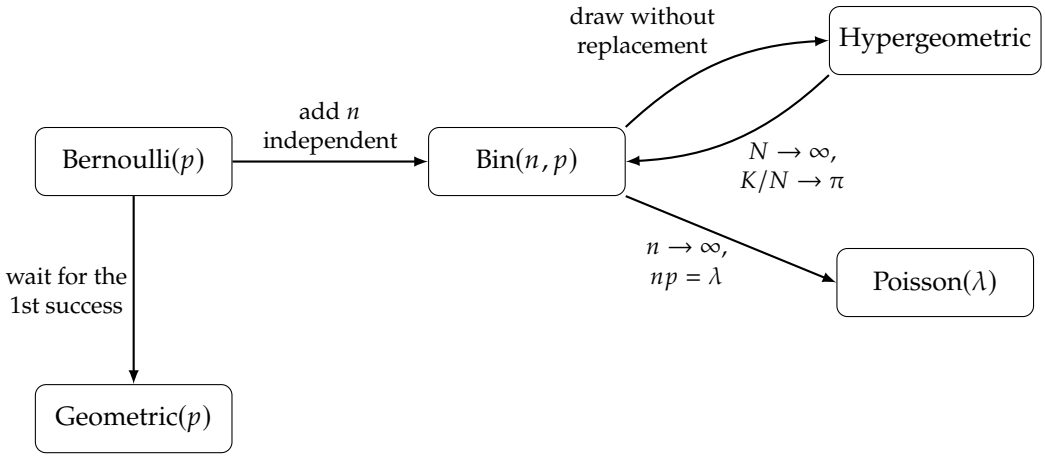


Figure 5: The five discrete models of this section. The arrow down from the Bernoulli changes what we count. The arrows out of the binomial change how we sample or take a limit. The two arrows between the binomial and the hypergeometric go both ways. Drawing without replacement turns the binomial into the hypergeometric. A large population turns the hypergeometric back into the binomial (the Remark in §5.4). The binomial becomes Poisson when trials are many and successes are rare (Theorem 5.21).

Is the number of trials fixed, or is the number of successes fixed?

A fixed number of trials and a random count of successes gives the binomial or the hypergeometric. A fixed number of successes, here one, and a random count of trials gives the geometric. This is the most useful distinction. Students often confuse the two cases.

Are the trials independent?

They are independent if we draw with replacement. They are almost independent if the population is very large. In both cases the binomial applies. Without replacement from a small population they are not, and the hypergeometric applies.

Are we counting events in a window instead of trials?

Then there are no trials to count, only a rate and a window, and the Poisson applies.

Two sanity checks are worth carrying into Section 6. First, the Poisson's mean and variance are equal. That is a strong restriction. In practice it is a common reason a Poisson model is rejected. Second, the hypergeometric's variance is never larger than the matching binomial's, and it is smaller whenever $n \geq 2$ and $0 < K < N$. Two facts give this. The finite-population correction is below 1 for $n \geq 2$. And $\pi(1 - \pi) > 0$ when $0 < K < N$. So a count drawn without replacement varies less than one drawn with replacement.

Exercises

Exercise 5.1. *$X \sim$ Bernoulli(0.4). Give the pmf, $\mathbb{E}(X)$, $\mathrm{Var}(X)$, and the MGF.*

Exercise 5.2. *$X \sim$ Bin$(8, 0.5)$. Find $P(X = 4)$, $\mathbb{E}(X)$, and $\mathrm{Var}(X)$.*

Exercise 5.3. *$X \sim$ Bin$(12, 0.2)$. Find $P(X = 2)$, $P(X \le 1)$, $\mathbb{E}(X)$, and $\mathrm{Var}(X)$.*

Exercise 5.4. *A quiz has 10 multiple-choice questions, each with 4 choices. A student guesses every answer at random, independently. Let X be the number of correct answers. Find $\mathbb{E}(X)$ and σ_X. The pass mark is 6 correct answers. Find the probability that the student passes.*

Exercise 5.5. *Let $X \sim$ Bin(n, p). Show that $n - X \sim$ Bin$(n, 1 - p)$, and say in one sentence what that means about Figure 1.*

Exercise 5.6. *Independent trials each succeed with probability $p = 0.2$; let X be the trial of the first success. Find $P(X = 1)$, $P(X = 3)$, $\mathbb{E}(X)$, and $P(X > 5)$.*

Exercise 5.7. *For the geometric variable of Exercise 5.6, verify memorylessness by computing $P(X > 7 \mid X > 4)$ and checking it equals $P(X > 3)$.*

Exercise 5.8. *A basketball player makes each free throw with probability 0.7, independently. Let X be the number of throws up to and including the first make. Find $P(X \le 2)$ and $\mathbb{E}(X)$. Then find the smallest x with $P(X \le x) \ge 0.99$.*

Exercise 5.9. *A batch of $N = 20$ items contains $K = 8$ defectives. Draw $n = 4$ without replacement; let X be the number of defectives drawn. Find the probability of at least one defective, $\mathbb{E}(X)$, and $\mathrm{Var}(X)$.*

Exercise 5.10. *An analyst picks 3 stocks at random, without replacement, from a list of 10. Exactly 4 of the 10 will beat the market this year. Let X be the number of winners among the 3 picks. Find $P(X \ge 2)$, $\mathbb{E}(X)$, and $\mathrm{Var}(X)$.*

Exercise 5.11. *A sample of $n = 10$ is drawn without replacement. Evaluate the finite-population correction $(N - n)/(N - 1)$ at $N = 20$, $N = 100$ and $N = 1000$, and say what each value means for the spread of the count.*

Exercise 5.12. *$X \sim$ Poisson(2). Find $P(X = 0)$, $P(X = 1)$, and $P(X \ge 2)$.*

Exercise 5.13. *Defects occur at $\lambda = 4$ per batch (Poisson). Find $P(X = 3)$, $\mathbb{E}(X)$, and $\mathrm{Var}(X)$.*

Exercise 5.14. *A Poisson variable X satisfies $P(X = 1) = P(X = 2)$. Find λ, then find $P(X = 0)$ and $P(X \ge 3)$.*

Exercise 5.15. *A production run has $n = 500$ chips, each failing independently with probability $p = 0.004$. Use a Poisson approximation to estimate P(at most one failure).*

Exercise 5.16. *$X \sim$ Poisson(1.5) and $Y \sim$ Poisson(2.5) are independent. Name the distribution of $X + Y$ and find $P(X + Y = 2)$.*

Exercise 5.17. *$X \sim$ Bin$(3, 0.5)$ and $Y \sim$ Bin$(4, 0.5)$ are independent. Use MGFs to show that $X + Y \sim$ Bin$(7, 0.5)$, and find $P(X + Y = 3)$.*

Exercise 5.18. *A lot of $N = 1000$ contains $K = 200$ defectives; draw $n = 5$ without replacement. Find the exact $P(X = 1)$ and the Bin$(5, 0.2)$ approximation, and say why they are so close here.*

Exercise 5.19. *Name the distribution each situation calls for, and say which of the three questions in §5.6 decided it. (a) Twelve independent machines each fail today with probability* 0.03*; count the failures. (b) Cards are dealt from a* 52*-card deck; count the aces in a hand of five. (c) A recruiter calls candidates, who each accept with probability* 0.2 *independently, until one accepts; count the calls. (d) Typing errors appear at an average rate of* 1.5 *per page; count the errors on one page.*

Exercise 5.20. *Let* $X \sim \text{Poisson}(\lambda)$*. Without using the MGF, show directly from the pmf that* $\mathbb{E}(X) = \lambda$ *and* $\mathbb{E}\big[X(X-1)\big] = \lambda^2$*. Then use these two results to find* $\text{Var}(X)$*.*

Exercise 5.21. *A stock price moves once a day. It rises by* 10 *with probability* 0.6 *and falls by* 10 *with probability* 0.4*. The moves on different days are independent. A trader buys one share at today's close. Call tomorrow day* 1*. The trader sells the share at the close of the first day on which the price falls. Let* T *be that day, so the trader holds the share for* T *days. (a) Find* $\mathbb{E}(T)$*. (b) Write the trader's profit in terms of* T*, and find the expected profit. (c) The profit can be positive, zero or negative. Find the probability of each case. (d) The price has risen on each of days* 1*,* 2 *and* 3*. Find the probability that it also rises on days* 4 *and* 5*.*

Exercise 5.22. *The number of defaults among a bank's loans in one month is* $\text{Poisson}(1)$*. The numbers of defaults in different months are independent. (a) Find the probability of no default in a month, and the probability of at least two defaults in a month. (b) Find the probability of at most one default in a quarter of three months. (c) Call a month* clean *if it has no default. Let* Y *be the number of clean months in a year of* 12 *months. Find* $\mathbb{E}(Y)$*. (d) Find the probability that the first clean month of the year is the third month. (e) Find the probability that a quarter has no default at all. Do it in two ways: from the count for the quarter, and from the three months one by one.*

Exercise 5.23. *A portfolio holds* 5 *bonds. Each bond defaults during the year with probability* 0.2*, independently of the others. At the end of the year a bond is worth* 100 *if it has not defaulted. It is worth* $100 - K$ *if it has defaulted, where* $0 < K \leq 100$*. Let* X *be the number of defaults. Let* V *be the value of the portfolio at the end of the year. (a) Write* V *in terms of* X *and* K*. (b) Take* $K = 40$*. Find* $P(V < 400)$*. (c) Find the largest* K *for which* $P(V < 400) \leq 0.10$*.*

Exercise 5.24. *An office files* 20 *expense claims in a month. Exactly* 4 *of them are false. An auditor checks* n *of the claims. The audit succeeds if it finds at least one false claim. (a) The auditor chooses* $n = 5$ *claims at random without replacement. Find the probability that the audit succeeds. (b) Suppose instead that the auditor draws with replacement. Each checked claim is put back, so it can be drawn again. Find the probability that the audit succeeds with* 5 *draws. (c) Go back to drawing without replacement. Find the smallest* n *for which the audit succeeds with probability at least* 0.9*. (d) Now draw with replacement. Find the smallest number of draws for which the audit succeeds with probability at least* 0.9*.*

Exercise 5.25. *A broker receives buy orders and sell orders. In one minute the number of buy orders is* $X \sim \text{Poisson}(2)$*. The number of sell orders is* $Y \sim \text{Poisson}(1)$*.* X *and* Y *are independent. Let* $S = X + Y$ *be the total number of orders in the minute. (a) Find* $P(S = 0)$ *and* $P(S = 3)$*. (b) Exactly three orders arrived in the minute. Find the probability that all three are buy orders. (c) Exactly three orders arrived in the minute. Find the probability that exactly two of them are buy orders. (d) Compare your answers to (b) and (c) with the pmf of* $\text{Bin}(3, \frac{2}{3})$*.*