Week 04 notes

Bernoulli process

01 Theory

In a Bernoulli process, an experiment with binary outcomes is repeated; for example flipping a coin repeatedly. Several discrete random variables may be defined in the context of some Bernoulli process.

Notice that the sample space of a Bernoulli process is infinite: an outcome is any sequence of trial outcomes, e.g. $HTHHTTHHHTTTHHHHTTTT\cdots$

Bernoulli variable

A random variable X_i is a **Bernoulli indicator**, written $X_i \sim \text{Ber}(p)$, when X_i indicates whether a success event, having probability p, took place in trial number i of a Bernoulli process.

Bernoulli indicator PMF:

$$P_{X_i}(1) = p$$

$$P_{X_i}(0) = 1 - p$$

$$P_{X_i}(x) = 0 \qquad (x \neq 1,\, 0)$$

- An RV that always gives either 0 or 1 for every outcome is called an **indicator** variable.

Binomial variable

A random variable S is **binomial**, written $S \sim \text{Bin}(n,\, p)$, when S counts the *number of successes* in a Bernoulli process, each having probability p, over a specified number n of trials.

Binomial PMF:

$$P_S(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

- For example, if $S \sim \text{Bin}(10,\, 0.2)$, then $P_S(5)$ gives the odds that success happens exactly 5 times over 10 trials, with probability 0.2 of success for each trial.
- In terms of the Bernoulli indicators, we have: $S = X_1 + X_2 + \cdots + X_n$
- If A is the success event, then $p = P[A]$, and $1 - p$ is the probability of failure.

Geometric variable

A random variable N is **geometric**, written $N \sim \text{Geom}(p)$, when N counts the *discrete wait time* until a *single* success takes place, given that success has probability p in each trial.

Geometric PMF:

$$P_N(k) = (1-p)^{k-1}p$$

- For example, if $N \sim \text{Geom}(30\%)$, then $P_N(7)$ gives the probability of getting: failure on the first 6 trials AND success on the 7^{th} trial.

Pascal variable

A random variable L is **Pascal**, written $L \sim \text{Pasc}(\ell, p)$, when L counts the *discrete wait time* until success has taken place *k times*, given that success has probability p in each trial.

Pascal PMF:

$$P_L(k) = \binom{k-1}{\ell-1}(1-p)^{k-\ell}p^{\ell}$$

- For example, if $L \sim \text{Pasc}(3, 0.25)$, then $P_L(8)$ gives the probability of getting: the 3^{rd} success on (precisely) the 8^{th} trial.
- Interpret the formula: # ways to arrange 2 successes among 7 'prior' trials, times the probability of exactly 3 successes and 5 failures in one particular sequence.
- The Pascal distribution is also called the **negative binomial** distribution, e.g. $L \sim \text{Negbin}(\ell, p)$.

Uniform variable

A discrete random variable X is **uniform** on a finite set $A \subset S$, written $X \sim \text{Unif(A)}$, when the probability is a fixed constant for outcomes in A and zero for outcomes outside A.

Discrete uniform PMF:

$$P_X(k) = \begin{cases} \frac{1}{|A|} & \text{when } k \in A \\ 0 & \text{when } k \notin A \end{cases}$$

Continuous uniform PDF:

$$f_X(x) = \begin{cases} \frac{1}{P[A]} & \text{when } x \in A \\ 0 & \text{when } x \notin A \end{cases}$$

02 Illustration

Example - Repeated die rolls, how many ones?

Let X have distribution given by this PMF:

x	1	2	3	4	5
$p_X(x)$	1/7	1/14	3/14	2/7	2/7

Find $E[|X-2|]$.

Solution

1. Compute the PMF.
   - PMF arranged by possible value:

$$\begin{array}{lll} P[|X-2|=0] & \gg\gg & P[X=2]=\frac{1}{14} \\ P[|X-2|=1] & \gg\gg & P[X=1]+P[X=3]=\frac{1}{7}+\frac{3}{14}=\frac{5}{14} \\ P[|X-2|=2] & \gg\gg & P[X=4]=\frac{2}{7} \\ P[|X-2|=3] & \gg\gg & P[X=5]=\frac{2}{7} \\ P[|X-2|=k] & \gg\gg & 0 \quad \text{for} \quad k \neq 0,1,2,3. \end{array}$$

2. Calculate the expectation.
   - Use discrete formula:

$$E[|X-2|]=0\cdot\tfrac{1}{14}+1\cdot\tfrac{5}{14}+2\cdot\tfrac{2}{7}+3\cdot\tfrac{2}{7}=\tfrac{25}{14}$$

Example - Roll die until

Roll a fair die repeatedly. Find the probabilities that:

- (a) At most 2 threes occur in the first 5 rolls
- (b) There is no three in the first 4 rolls
- (c) There is no three in the first 4 rolls, but there is at least one in the first 20 rolls

Solution

(a)

1. Labels.
   - Use $S_5 \sim \text{Bin}(6, 1/6)$ to count the number of threes among the first six rolls.
   - Seek $P[S_5 \leq 2]$ as the answer.
2. Calculations.
   - Divide into exclusive events:

$$P[S_5 \leq 2] \quad \gg\gg \quad P[S_5=0]+P[S_5=1]+P[S_5=2]$$

$$\gg\gg \quad \binom{5}{0}\left(\frac{1}{6}\right)^0\left(\frac{5}{6}\right)^5+\binom{5}{1}\left(\frac{1}{6}\right)^1\left(\frac{5}{6}\right)^4+\binom{5}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^3$$

$$\gg\gg \quad \frac{625}{648}$$

(b)

1. Labels.
   - Use $N \sim \text{Geom}(1/6)$ to give the roll number of the first time a three is rolled.
   - Seek $P[N > 4]$ as the answer.
2. Sum the PMF formula for Geom(1/6).
   - Compute:

$$P[N>4] \quad \gg\gg \quad \sum_{k=5}^{\infty}\left(\frac{5}{6}\right)^{k-1}\left(\frac{1}{6}\right)$$

3. Geometric series formula.

- For any geometric series:

$$a + ar + ar^2 + ar^3 + \cdots \quad \gg\gg \quad \frac{a}{1-r}$$

- Apply formula:

$$\sum_{k=5}^{\infty} \left(\frac{5}{6}\right)^{k-1} \left(\frac{1}{6}\right) \quad \gg\gg \quad \left(\frac{5}{6}\right)^4$$

4. Final answer is $P[N > 4] = (5/6)^4$.

(c)

1. Labels.
   - Event A means "no three in the first 4 rolls."
   - Event B means "no three in rolls 5 through 20."
   - These events are independent!
   - Seek $P(AB^c)$ as answer.
2. Known probabilities.
   - Know: $P[A] = (5/6)^4$
   - Know: $P[B] = (5/6)^{16}$
3. Apply product rule for independent events.
   - Product rule and negation rule:

$$P(AB^c) \quad \gg\gg \quad P(A)(1 - P(B))$$

   - Insert data:

$$\gg\gg \quad \left(\frac{5}{6}\right)^4 \left(1 - \left(\frac{5}{6}\right)^{16}\right) \quad \gg\gg \quad \approx 0.456$$

Expectation and variance

03 Theory

Expected value

The **expected value** $E[X]$ of random variable X is the weighted average of the values of X, weighted by the probability of those values.

Discrete formula using PMF:

$$E[X] = \sum_k k \cdot P_X(k)$$

Continuous formula using PDF:

$$E[X] = \int_{-\infty}^{+\infty} x \cdot f_X(x)\, dx$$

Notes:

- Expected value is sometimes called **expectation**, or even just **mean**, although the latter is best reserved for statistics.

- The Greek letter μ is also used in contexts where 'mean' is used.

Let X be a random variable, and write $E[X] = \mu$.

Variance

The **variance** $\text{Var}[X]$ measures the average *squared deviation* of X from μ. It estimates how *concentrated* X is around μ.

- Defining formula:

$$\text{Var}[X] = E[(X - \mu)^2]$$

- Sometimes easier formula:

$$\text{Var}[X] = E[X^2] - E[X]^2$$

Calculating variance

- Discrete formula using PMF:

$$\text{Var}[X] = \sum_k (k - \mu)^2 P_X(k)$$

- Continuous formula using PDF:

$$\text{Var}[X] = \int_{-\infty}^{+\infty} (x - \mu)^2 f_X(x)\, dx$$

Standard deviation

The quantity $\sigma_X = \sqrt{\text{Var}[X]}$ is called the **standard deviation** of X.

04 Illustration

Exercise - Tokens in bins

Consider a game like this: a coin is flipped; if H then draw a token from Bin 1, if T then from Bin 2.

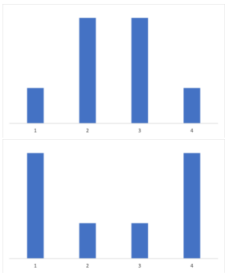
- Bin 1 contents: 1 token \$1,000, and 9 tokens \$1
- Bin 2 contents: 5 tokens \$50, and 5 tokens \$1

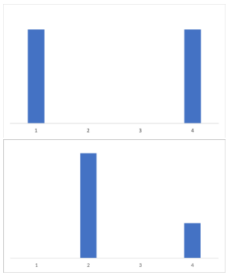
It costs \$50 to enter the game. Should you play it? (A lot of times?) How much would you pay to play?

Solution

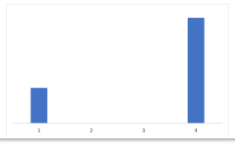
Expectations

Same, same, same...





... different



Example - Expected value: rolling dice

Let X be a random variable counting the number of dots given by rolling a single die.

Then:

$$E[X] \quad \gg\gg \quad 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + \cdots + 6 \cdot \frac{1}{6} \quad \gg\gg \quad \frac{7}{2}$$

Let S be an RV that counts the dots on a roll of *two* dice.

The PMF of S:

k	2	3	4	5	6	7	8	9	10	11	12
$p_S(k) = P(S = k)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Then:

$$E[S] \quad \gg\gg \quad 2 \cdot \frac{1}{36} + 3 \cdot \frac{2}{36} + 4 \cdot \frac{3}{36} + \cdots + 12 \cdot \frac{1}{36} \quad \gg\gg \quad 7$$

- Notice that $\frac{7}{2} + \frac{7}{2} = 7$.
  - In general, $E[X + Y] = E[X] + E[Y]$.
  - Let X be a green die and Y a red die.
  - From the earlier calculation, $E[X] = \frac{7}{2}$ and $E[Y] = \frac{7}{2}$.
  - Since $S = X + Y$, we derive $E[S] = 7$ by simple addition!

Example - Expected value by finding new PMF

Let X have distribution given by this PMF:

x	1	2	3	4	5
$p_X(x)$	1/7	1/14	3/14	2/7	2/7

Find $E[|X - 2|]$.

Solution

1. Compute the PMF.

- PMF arranged by possible value:

$$\begin{array}{lll} P[|X-2|=0] & \gg\gg & P[X=2]=\frac{1}{14} \\ P[|X-2|=1] & \gg\gg & P[X=1]+P[X=3]=\frac{1}{7}+\frac{3}{14}=\frac{5}{14} \\ P[|X-2|=2] & \gg\gg & P[X=4]=\frac{2}{7} \\ P[|X-2|=3] & \gg\gg & P[X=5]=\frac{2}{7} \\ P[|X-2|=k] & \gg\gg & 0 \quad \text{for} \quad k \neq 0,1,2,3. \end{array}$$

2. ☰ Calculate the expectation.
   - Use discrete formula:

$$E[|X-2|]=0\cdot\tfrac{1}{14}+1\cdot\tfrac{5}{14}+2\cdot\tfrac{2}{7}+3\cdot\tfrac{2}{7}=\tfrac{25}{14}$$

Exercise - Variance using simplified formula

Suppose X has this PMF:

k :	**1**	**2**	**3**
$P_X(k)$:	1/7	2/7	4/7

Find $\text{Var}[\frac{1}{1+X}]$ using the formula $\text{Var}[Y]=E[Y^2]-E[Y]^2$ with $Y=\frac{1}{1+X}$.

Partial answer: $E[Y]=\frac{13}{42}$ and $E[Y^2]=\frac{13}{126}$.

Poisson process

05 Theory

Poisson variable

A random variable X is **Poisson**, written $X \sim \text{Pois}(\lambda)$, when X counts the number of "arrivals" in a fixed "interval." It is applicable when:

- The arrivals come at a *constant average rate* λ.
- The arrivals are independent of each other.

Poisson PMF:

$$P_X(k)=e^{-\lambda}\frac{\lambda^k}{k!}$$

A Poisson variable is comparable with a binomial variable. Both count the occurrences of some "arrivals" over some "space of opportunity."

- The binomial opportunity is a set of *n repetitions* of a trial.
- The Poisson opportunity is a *continuous interval* of time.

In the binomial case, success occurs at some rate p, since $p=P[A]$ where A is the success event.

In the Poisson case, arrivals occur at some rate λ.

The Poisson distribution is actually the limit of binomial distributions by taking $n \to \infty$ while np remains fixed, so $p \to 0$ in perfect balance with $n \to \infty$.

Let $X_{n,p} \sim \text{Bin}(n, p)$ and let $Y_\lambda \sim \text{Pois}(\lambda)$. Fix λ and let $p = \lambda/n$. Then for any $k \in \mathbb{N}$:

$$P_{X_{n,p}}(k) \quad \overset{n\to\infty}{\longrightarrow} \quad P_{Y_\lambda}(k)$$

Interpretation - Binomial model of rare events

Let us interpret the assumptions of this limit. For n large but p small such that $\lambda = np$ remains moderate, the binomial distribution describes a large number of trials, a low probability of success per trial, but a moderate total count of successes.

This setup describes a physical system with a large number of parts that may activate, but each part is unlikely to activate; and yet the number of parts is so large that the total number of arrivals is still moderate.

06 Illustration

Example - Radioactive decay is Poisson

Consider a macroscopic sample of Uranium.

Each atom decays independently of the others, and the likelihood of a single atom popping off is very low; but the product of this likelihood by the total number of atoms is a moderate number.

So there is some constant average rate of atoms in the sample popping off, and the number of pops per minute follows a Poisson distribution.

Calls to a call center is Poisson

Consider a call center that receives help requests from users of a popular phone manufacturer.

The total number of users is very large, and the likelihood of any given user calling in a given minute is very small, but the product of these rates is moderate.

So there is some constant average rate of calls to the center, and the number of calls per minute follows a Poisson distribution.

Exercise - Typos per page

A draft of a textbook has an average of 6 typos per page.

What is the probability that a randomly chosen page has ≥ 4 typos?

Answer: 0.849

Hint: study the complementary event.

Exercise - Poisson calculation

Suppose $X \sim \text{Pois}(10)$. Find $P(X \leq 13 \mid X \geq 7)$.

07 Theory

Extra- Derivation of binomial limit to Poisson

Consider a random variable $X \sim \text{Bin}(n, p)$, and suppose n is very large.

Suppose also that p is very small, such that $E[X] = np$ is *not* very large, but the extremes of n and p counteract each other. (Notice that then npq will *not* be large so the normal approximation does *not* apply.) In this case, the binomial PMF can be approximated using a factor of e^{-np}. Consider the following rearrangement of the binomial PMF:

$$\begin{aligned} p_X(k) &= \binom{n}{k} p^k q^{n-k} \\ &= \frac{n(n-1)\cdots(n-k+1)}{k!} p^k (1-p)^n \frac{1}{q^k} \\ &= (1-p)^n \frac{(np)^k}{k!} \left[\frac{n}{n} \cdot \frac{n-1}{n} \cdot \frac{n-2}{n} \cdots \frac{n-k+1}{n}\right] \frac{1}{q^k}. \end{aligned}$$

Since n is very large, the factor in brackets is approximately 1, and since p is very small, the last factor of $1/q^k$ is also approximately 1 (provided we consider k small compared to n). So we have:

$$p_X(k) \approx (1-p)^n \frac{(np)^k}{k!}.$$

Write $\lambda = np$, a moderate number, to find:

$$p_X(k) \approx \left(1 - \frac{\lambda}{n}\right)^n \frac{\lambda^k}{k!}.$$

Here at last we find $e^{-\lambda}$, since $\left(1 - \frac{\lambda}{n}\right)^n \to e^{-\lambda}$ as $n \to \infty$. So as $n \to \infty$:

$$p_X(k) \approx e^{-\lambda} \frac{\lambda^k}{k!}.$$

Extra - Binomial limit to Poisson and divisibility

Consider a sequence of increasing n with decreasing p such that $\lambda = np$ is held fixed. For example, let $n = 1, 2, 3, \ldots$ while $p = \frac{\lambda}{n}$.

Think of this process as increasing the number of causal agents represented: group the agents together into n bunches, and consider the odds that such a bunch activates. (For the call center, a bunch is a group of users; for radioactive decay, a bunch is a unit of mass of Uranium atoms.)

As n doubles, we imagine the number of agents per bunch to drop by half. (For the call center, we divide a group in half, so twice as many groups but half the odds of one group making a call; for the Uranium, we divide a chunk of mass in half, getting twice as many portions with half the odds of a decay occurring in each portion.

This process is formally called *division of a distribution*, and the fact that the Poisson distribution arises as the limit of such division means that it is infinitely divisible.

Extra - Theorem: Poisson approximation of the binomial

Suppose $X \sim \operatorname{Bin}(n,\ p)$ and $Y \sim \operatorname{Pois}(np)$. Then:

$$\left|P_X(k) - P_Y(k)\right| \leq np^2$$

for any $k \in \mathbb{N}$.

In consequence of this theorem, a Poisson distribution may be used to approximate the probabilities of a binomial distribution for large n when it is impracticable (even for a computer) to calculate large binomial coefficients.

The theorem shows that the Poisson approximation is appropriate when np is a moderate number while np^2 is a small number.