VE401 Recitation Class Note2

Discrete Random Variables

Chen Siyi
siyi.chen_chicy@sjtu.edu.cn

Idea of random variables: take elements of a sample space and map them into a subset of the real numbers, i.e.,

$$X : S \to \mathbb{R}$$

X has numerical values (“variable”) that are derived from the outcome of a random experiment (“random”).

We first focus on discrete random variables, and deal with continuous random variables later.

1 Overview

1.1 Definiation

Let S be a sample space and Ω a countable subset of R. A discrete random variable is a map

$$X : S \to \Omega$$

together with a function

$$f_X : \Omega \to \mathbb{R}$$

where

(i) $f_X(x) \geq 0$ for all $x \in \Omega$

(ii) $\sum_{x \in \Omega} f_X(x) = 1$

We often say that a random variable is given by the pair (X , f_X).

1.2 General Properties

Various distributions’ important features:

1. $f_X = P[X = x]$: probability density function(PDF).
2. $F_X(x) = \sum_{y \leq x} f_X(y)$: cumulative distribution function(CDF).
3. $\mathrm{E}[X] := \sum_{x \in \Omega} x \cdot f_X(x)$: expectation.
4. $\mathrm{Var}[X] := \mathrm{E}\left[(X - \mathrm{E}[X])^2\right] = E[X^2] - E[X]^2$: Variance.
5. $m_X(t) := \sum_{k=0}^{\infty} \frac{E[X^k]}{k!} t^k = E[e^{tX}]$: moment generating function (MGF).

2 Expectation

Definition:

$$\mathrm{E}[X] := \sum_{x\in\Omega} x \cdot f_X(x)$$

Exists only if E[X] converges.

Properties:

1. $\mathrm{E}[\varphi \circ X] = \sum_{x\in\Omega} \varphi(x) \cdot f_X(x)$
2. $c \in \mathbb{R}$, then E[c]=c, E[cX]=cE[X]
3. X, Y both be random variables, then E[X+Y]=E[X]+E[Y]

Comments:

1. Describe the location of the average value
2. Different from "median"(or modes)

Question1: Two Envelopes Problem

You are given two indistinguishable envelopes, each containing money, one contains twice as much as the other. You may pick one envelope and keep the money it contains. Having chosen an envelope at will, but before inspecting it, you are given the chance to switch envelopes. Should you switch?
A student gives a solution:
Suppose in the chosen envelop, there's M amount of money.
Then the expectation of the other envelope's money is:

$$\frac{1}{2} \cdot 2M + \frac{1}{2} \cdot 0.5M = 1.25M > M$$

So we should switch. Is this right?

3 Variance

Definition:

$$\mathrm{Var}[X] := \mathrm{E}\left[(X - \mathrm{E}[X])^2\right] = E[X^2] - E[X]^2$$

Standard deviation:

$$\sigma_X = \sqrt{Var[X]}$$

Properties:

1. $c \in \mathbb{R}$, then Var[c]=0, Var[cX]=c^2Var[X]
2. Var[X+Y]? See later. In general, for dependent X and Y, $Var[X+Y] \neq Var[X]+Var[Y]$

4 Standardize Random Variables

Definition:

X is a given random variable with E[X]=μ, Var[X]=σ^2, then the standardized variable is:

$$Y = \frac{X-\mu}{\sigma}$$

where

$$E[Y] = 0$$

$$Var[Y] = 1$$

Purpose:

To help make reasonable comparisons. Will see more clearly when performing statistic tests later.

5 Moment Generating Functions

n_{th} (ordinary) moments of X: E[X^n], n=1,2,3,...

n_{th} central moments of X: E$\left[(\frac{X-\mu}{\sigma})\right]$, n=3,4,5,...

Moment generating function:

$$m_X(t) := \sum_{k=0}^{\infty} \frac{E\left[X^k\right]}{k!} t^k = E[e^{tX}]$$

having radius of convergence $\varepsilon > 0$. Application:

$$\mathrm{E}\left[X^k\right] = \left.\frac{d^k m_X(t)}{dt^k}\right|_{t=0}$$

Example1: Find E for Geom

$m_X(t) = E[e^{tX}] = \sum_{x=1}^{\infty} e^{tx} q^{x-1} p = \frac{p}{q} \sum_{x=1}^{\infty} (qe^t)^x = \frac{pe^t}{1-qe^t}$

$E[X] = \frac{d}{dt}\Big|_{t=0} m_X(t) = \frac{d}{dt}\Big|_{t=0} \frac{p}{e^{-t}-q} = \frac{1}{p}$

6 Independent and Identical

Independent: the outcome of one trial does not influence the outcome of the following trials.

Identical: each trial has the same probability of success.

Example2: Dependent but Identical

Suppose a box is filled with 10 red balls and 10 black balls. You pick out 20 balls one by one without putting back. All the 20 trials are identical but not independent. Why?

1. The results of the previous draws apparently affect the following ones. So dependent.
2. Without taking any action(without knowing the color of any drawn balls), the probability for any ball to be red is the same, equals to 0.5. So identical. Understand: Model the process of drawing 20 balls out as order the 20 balls to a sequence. The position does not affect the probability of the ball at that position being red.

7 Specific Distributions

7.1 Bernoulli Distribution

$$X \sim Bernoulli(p)$$

Interpolation:

Perform one trial. Only two possible outcomes. Probability for success is p, for failure is q=1-p. x=1 means success, and x=0 means failure.

Features:

1. p is the parameter
2. $f_X(x) = \begin{cases} 1-p & \text{for } x=0 \\ p & \text{for } x=1 \end{cases}$
3. E[X]=p
4. Var[X]=pq

7.2 Binomial Distribution

$$X \sim B(n,p)$$

Interpolation:

Perform n independent and identical Bernoulli trials with parameter p. X gives the total number of success in n trials.

Features:

1. p, n are the parameters
2. $f_X(x) = \left(\begin{array}{c} n \\ x \end{array} \right) p^x(1-p)^{n-x}$

3. $F_X(x) = \sum_{y=0}^{\lfloor x \rfloor} \begin{pmatrix} n \\ y \end{pmatrix} p^y(1-p)^{n-y}$

4. E[X]=np

5. Var[X]=npq

6. $m_X(t) = (q+pe^t)^n, \quad m_X : \mathbb{R} \to \mathbb{R}$

Question2: Banach Matchbox Problem

Suppose a mathematician carries two matchboxes at all times: one in his left pocket and one in his right. Each time he needs a match, he is equally likely to take it from either pocket. Suppose he reaches into his pocket and discovers for the first time that the box picked is empty and he failed to take one out. If it is assumed that each of the matchboxes originally contained n matches, what is the probability that there are exactly k matches in the other box?

A student gives a solution:

Event X: an empty box and a x-matches box are left.

Ways of X: $\begin{pmatrix} 2n-x \\ n \end{pmatrix} \cdot 2$

Total ways:

By Cardano's Principle,

$$P[x=k] = \frac{\begin{pmatrix} 2n-k \\ n \end{pmatrix} \cdot 2}{\sum_{i=0}^{n} \begin{pmatrix} 2n-i \\ n \end{pmatrix} \cdot 2} = \frac{\begin{pmatrix} 2n-k \\ n \end{pmatrix}}{\sum_{i=0}^{n} \begin{pmatrix} 2n-i \\ n \end{pmatrix}}$$

Is this right?

7.3 Geometric Distribution

$$X \sim Geom(p)$$

Interpolation:

Perform a sequence of i.i.d. Bernoulli trials with parameter p, and stop until get a success. X gives the total number of trials needed to obtain the first success.

Features:

1. p is the parameter

2. $f_X(x) = (1-p)^{(x-1)}p$

3. F(x)=$1 - q^{\lfloor x \rfloor}$

4. E[X]=$\frac{1}{p}$
5. Var[X]=$\frac{q}{p^2}$
6. $m_X(t) = \frac{pe^t}{1-qe^t}, \quad m_X : (-\infty, -\ln q) \to \mathbb{R}$

7.4 Pascal Distribution

Interpolation:

Perform a sequence of i.i.d. Bernoulli trials with parameter p, and get the r^{th} success at the x^{th} trial.

Features:

1. p, r are the parameters
2. $f_X(x) = \begin{pmatrix} x-1 \\ r-1 \end{pmatrix} p^r(1-p)^{x-r}$
3. E[X]=$\frac{r}{p}$
4. Var[X]=$\frac{rq}{p^2}$
5. $m_X(t) = \left(\frac{pe^t}{1-qe^t}\right)^r, \quad m_X : (-\infty, -\ln q) \to \mathbb{R}$

Comments:

1. The Pascal distribution is a generalization of the Geometric distribution. Stop until get r success.
2. A random variable following the Pascal distribution with parameters r and p is the sum of r independent geometric random variables with parameter p.

7.5 Negative Binomial Distribution

Interpolation:

Perform a sequence of i.i.d. Bernoulli trials with parameter p, and get the r^{th} success after x failures. The same as get the r^{th} success at the $(x+r)^{th}$ trial.

Features:

1. p, r are the parameters
2. $f_X(x) = \begin{pmatrix} x+r-1 \\ r-1 \end{pmatrix} p^r(1-p)^x = \begin{pmatrix} -r \\ x \end{pmatrix} (-1)^x p^r (1-p)^x$
3. E[X]=$\frac{r(1-p)}{p}$. Can you imagine why?

The Negative Binomial:

$$\begin{pmatrix} -r \\ x \end{pmatrix} = \begin{pmatrix} r-1+x \\ r-1 \end{pmatrix} (-1)^x$$

Question3: Pascal/Negative Binomial

A pediatrician wishes to recruit 5 couples to participate in a new natural childbirth regimen. Let p = P(a randomly selected couple agrees to participate) = 0.2. What is the probability that 15 couples must be asked before 5 are found who agree to participate?

7.6 Poisson Distribution

Interpolation:

In a continuous interval [a, b], a certain event occurs for totally x times.

Assumptions:

1. Independence: If the intervals T_1, $T_2 \subset$ [0, t] do not overlap (except perhaps at one point), then the numbers of arrivals in these intervals are independent of each other.
2. Constant rate of arrivals.

Features:

1. k is the parameter. $k = \lambda t$, where λ is the arrival rate and t is the length of the interval [a, b].
2. $f_X(x) = \frac{k^x e^{-k}}{x!}$, x = 0, 1, 2, 3, ...
3. E[X]=k
4. Var[X]=k
5. $m_X(t) = e^{k(e^t - 1)}$

Question4: The "Discrete" Poisson

Which following property does the PDF of the Poisson distribution satisfy?
Prove your choice.

$$\sum_{x \in \Omega} f_X(x) = 1$$

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

Comments:

1. The Poisson random variable is a discrete random variable just in a continuous environment. Although its probability function $f_X(x)$ is a continuous function, the value of the random variable X are discrete points.

2. How to interpolate ”k”? When you consider the number of arrivals in an interval [a, b], k is the **expected total** number of arrivals in [a, b].

3. **A binomial distribution with large n and small p, can be approximated by a Poisson distribution with k=np**, where you make sure the expected value is the same. Because when $n \to \infty, n \cdot p = k$:

$$\begin{pmatrix} n \\ x \end{pmatrix} p^x(1-p)^{n-x} = \frac{k^x}{x!}e^{-k}$$

4. E[X]=Var[X]=k is an important and useful conclusion.

Question5: The Restaurant Problem

Suppose a fast food restaurant can expect 2 patrons every 3 minutes on average. Now you know from 4:00 p.m to 4:06 p.m there are totally 5 patrons enter to this restaurant. Assume Poisson Distribution. What is the probability that 3 or fewer patrons enter the restaurant from 4:03 p.m to 4:06 p.m?
Student A gives a solution:
”We can just simply calculate within a period of 3 mins.”

$$P = \sum_{i=0,1,2,3} \frac{k^i e^{-k}}{i!} = \sum_{i=0,1,2,3} \frac{2^i e^{-2}}{i!}$$

Student B gives another one:
”We should also consider the given situation that 5 patrons enters within the total 6 mins.”
He set $k_1 = 2$, and $k_2 = 4$.

$$P = \left(\sum_{i=0,1,2,3} \frac{{k_1}^i e^{-k_1}}{i!} \right) \div \frac{{k_2}^5 e^{-k_2}}{5!}$$

Which one is right?

Question6: Poisson & Binomial

A computer terminal can pick up an erroneous signal from the keyboard that does not show up on the screen. This creates a silent error that is difficult to detect. Assume that, for a particular keyboard, the probability that the silent error will occur per entry is 1/1000. In 12,000 entries estimate the probability that exactly 5 silent errors occurs.

Question7: Poisson & Negative Binomial

Now you are familiar with approximating the Binomial distribution with the Poisson distribution. Notice the main point to interpolate the approximation is: when n is large enough, the discrete environment of the Binomial distribution can be approximated to a continuous one.

By the way, for a Negative Binomial distribution with parameters p and r, can we approximate it with a Poisson distribution? If we can, how to decide the parameter k for the Poisson distribution?

7.7 *Hypergeometric Distribution

8 Answers

Answer1: Two Envelopes Problem

Wrong. In the two cases, the total amount of money in the two envelopes are 3M and 1.5M. But the actual background have fixed the total amount of money to be the same.
True solution:
Suppose the total amount of money in two envelopes are M.
Then the expectation of the other envelope's money is:

$$\frac{1}{2}\cdot\frac{2}{3}M+\frac{1}{2}\cdot\frac{1}{3}M=\frac{1}{2}M$$

Same for the chosen envelop. So there's no need to switch.

Answer2: Banach Matchbox Problem

Wrong. Not every way has the same probability. Recall D'Alembert's Error.
True solution:
Consider in the first 2n-k draws between two box A and B, choose n draws of A and n-k draws of B. Then in the 2n-k+1 draw, also choose A. Similarly, exchange A and B. Therefore:

$$P[x=k]=\begin{pmatrix} 2n-k \\ n \end{pmatrix}\left(\frac{1}{2}\right)^{n}\left(\frac{1}{2}\right)^{n-k}\frac{1}{2}\times 2=\begin{pmatrix} 2n-k \\ n \end{pmatrix}\left(\frac{1}{2}\right)^{2n-k}$$

Answer3: Pascal/Negative Binomial

Very easy to go back to a simple model.
Pascal: Get the 5^{th} success at the 15^{th} trial. Pascal distribution with p = 0.2, r = 5, x = 15.

$$P[x = 15] = \left(\begin{array}{c} 14 \\ 4 \end{array} \right) (0.2)^5 (0.8)^{10} = 0.034$$

Negative Binomial: Get the 5^{th} success after 10 failures. Negative Binomial distribution with p = 0.2, r = 5, x = 10.

$$P[x = 15] = \left(\begin{array}{c} 14 \\ 4 \end{array} \right) (0.2)^5 (0.8)^{10} = 0.034$$

You can say they are actually the "same" model, which also has close relationship with the Binomial distribution model.

Answer4: The "Discrete" Poisson

The Poisson random variable is a discrete random variable. You should be confidence enough to say the right answer is:

$$\sum_{x \in \Omega} f_X(x) = 1$$

The proof is also simple. Just recall the Maclaurin series of e^k:

$$e^k = 1 + k + \frac{k^2}{2!} + \frac{k^3}{3!} + \frac{k^4}{4!} + ...$$

Then:

$$\sum_{x \in \Omega} f_X(x) = \sum_{x=0}^{\infty} \frac{e^{-k} k^x}{x!} = e^{-k} e^k = 1$$

This emphasizes the "discrete" property of the Poisson distribution.

Answer5: The Restaurant Problem

This is a problem mixing the Poisson Distribution with the conditional probability. Be careful.
First, we figure out what we should solve:
Event A: 5 patrons enter within 4:00 p.m to 4:06 p.m;
Event B: 3 or fewer patrons enter within 4:03 p.m to 4:06 p.m;
Event C: 3 or fewer patrons enter within 4:03 p.m to 4:06 p.m given 5 patrons enter within 4:00 p.m to 4:06 p.m;

$$P[C] = P[B|A] = \frac{P[A \cap B]}{P[A]}$$

Next, we find out P[A] and P[B|A] using a distribution model. Of course, here the model is the Poisson distribution. So how do we find parameter k for it? As mentioned before, the key is the "expected value".
For P[A], from 4:00 p.m to 4:06 p.m, totally 6 mins; so we expected 4 patrons, which means we set $k_1 = 4$, and calculate:

$$P[A] = \frac{e^{-k_1} {k_1}^x}{x!} = \frac{e^{-4} 4^5}{5!} = 0.1563$$

For P[$A \cap B$], be careful again. It actually means: i patrons enter within 4:03 p.m to 4:06 p.m, and (5-i) patrons enter within 4:00 p.m to 4:03 p.m, where i can only be 0, 1, 2, 3. For both 3-minute period, the expected number of patrons is 2, so we set $k_2 = 2$, and calculate:

$$P[A \cap B] = \sum_{i=0,1,2,3} \frac{e^{-k_2} {k_2}^i}{i!} \cdot \frac{e^{-k_2} {k_2}^{(5-i)}}{(5-i)!} = e^{-4} 2^5 \sum_{i=0,1,2,3} \frac{1}{i!(5-i)!} = 0.1270$$

Therefore:

$$P[C] = \frac{0.1270}{0.1563} = 81.25\%$$

Answer6: Poisson & Binomial

It's easy to see the problem initially follows a Binomial distribution with n= 12,000 being large and p=0.001 being small. So we approximate it with a Poisson distribution. Also based on the key point "the expected value", k = np = 12. So for x = 5:

$$P[X=5] = \frac{e^{-k_1}{k_1}^x}{x!} = \frac{e^{-12}12^5}{5!} = 1.27\%$$

Answer7: Poisson & Negative Binomial

You should also be confidence enough to answer: yes, we can. Then, how?
Before go to detailed mathematics, first let's imagine.

1. What does the Negative Binomial model means? What are the meanings of p, r, x?
2. Imagine you the Negative Binomial model with r large enough, what will you expect x to be?
3. And similarly, k can be set as what "x is expected to be".

Then let's see the math details. As the above assumption, we set $k = \frac{p}{1-p} \cdot r$, so $p = \frac{k}{k+r}$. Then when $r \to \infty$:

$$\begin{aligned} f_X(x) &= \frac{(x+r-1)(x+r-2)\cdots(r)}{x!} p^x (1-p)^r \\ &= \frac{(x+r-1)(x+r-2)\cdots(r)}{x!} \frac{k^x}{(k+r)^x} \left(\frac{r}{k+r}\right)^r \\ &= \frac{k^x}{x!} \left(\frac{1}{(1+\frac{k}{r})^r}\right) \\ &= \frac{k^x}{x!} e^{-k} \end{aligned}$$

You can also think about and try with the Pascal distribution. Since as we discussed before, Pascal and Negative Binomial can just be viewed as the same model.