VE401 Probabilistic Methods in Eng.
Recitation Class (Mid Part2)

TA:HAN Liying, SJTU-UMJI
Email:hanliying_17@sjtu.edu.cn

March 29, 2020

Discrete Random Variable

1 Fundamentals

1.1 Probability Density Function

A discrete random variable is a map $X : S \to \Omega$ together with a function $f_X : \Omega \to \mathbb{R}$ having the properties that

1. $f_X(x) \geq 0$ for all $x \in \Omega$ and
2. $\sum_{x\in\Omega} f_X(x) = 1$

Notation:
For discrete random variables, we define the density function f_X in such a way that

$$f_X(x) = P[X = x]$$

1.2 Cumulative Distribution Function

The cumulative distribution function of a random variable, defined as follows,

$$F_X(x) := P[X \leq x]$$

For a discrete random variable

$$F_X(x) = \sum_{y \leq x} f_X(y)$$

1.3 Expectation

Let (X, f_X) be a discrete random variable. Then the expected value of X is

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

provided that the sum (possibly series if Ω is infinite) on the right **converges absolutely**.

- St. Petersburg Paradox – Expectation doesn't exit.

Properties of the Expectation:

1. Let X be a random variable, and for some $c \in \mathbb{R}$,

$$\mathrm{E}[c] = c, \qquad \mathrm{E}[cX] = c\mathrm{E}[X]$$

2. Let X and Y be random variables,

$$\mathrm{E}[X+Y] = \mathrm{E}[X] + \mathrm{E}[Y]$$

3. Let (X, f_X) be a discrete random variable and $\varphi : \Omega \to \mathbb{R}$ some function. Then the expected value of $\varphi \circ X$ is

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

1.4 Variance

The variance of X is defined as the mean square deviation from the mean,

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

- Standard deviation: $\sigma_X = \sqrt{\operatorname{Var} X} = \sqrt{\sigma_X^2}$
- Using the properties of the mean, we can find that

$$\begin{aligned}
\operatorname{Var} X &= \mathrm{E}\left[(X - \mathrm{E}[X])^2\right] \\
&= \mathrm{E}\left[X^2 - 2\mathrm{E}[X] \cdot X + \mathrm{E}[X]^2\right] \\
&= \mathrm{E}\left[X^2\right] - \mathrm{E}[X]^2
\end{aligned}$$

Properties of the Variance
Let X be a random variable, and for some $c \in \mathbb{R}$,

$$\operatorname{Var} c = 0, \quad \operatorname{Var} cX = c^2 \operatorname{Var}[X]$$

1.5 Moments

n^{th} **(Ordinary) Moments of** X**:**

$$\mathrm{E}\left[X^n\right], \quad n \in \mathbb{N}$$

n^{th} **Central Moments of** X**:**

$$\mathrm{E}\left[\left(\frac{X-\mu}{\sigma}\right)^n\right], \quad n = 3, 4, 5, \ldots$$

Moment-generating function of X**:**
For (X, f_X) such that the sequence of $E\left[X^k\right]$, $n \in \mathbb{N}$ exists. If the power series

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

has radius of convergence $\varepsilon > 0$, the thereby the function

$$m_X(t) : (-\varepsilon, \varepsilon) \to \mathbb{R}$$

is called the m.g.f of X.

Theorem *Let $\varepsilon > 0$ be given such that $E\left[e^{tX}\right]$ exists and has a power series expansion in t that converges for $|t| < \varepsilon$. Then the moment-generating function exists and*

$$m_X(t) = \mathrm{E}\left[e^{tX}\right]$$

for $|t| < \varepsilon$ Furthermore,

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

2 Common Distributions

2.1 Bernoulli Distribution

Suppose that the probability of success is p, where $0 < p < 1$.

$$f_X(x) = \begin{cases} 1-p & \text{for } x = 0 \\ p & \text{for } x = 1 \end{cases}$$

Such an experiment is said to be a Bernoulli trial.

Example: an experiment that can result in two possible outcomes, like tossing a coin.
independent and identical trials

2.2 Binomial Distribution

Meaning: the probability of getting x successes in n Bernoulli trials.

$$f_X(x) = P[x \text{ successes in } n \text{ trials }] = \begin{pmatrix} n \\ x \end{pmatrix} p^x (1-p)^{n-x}$$

Example: getting 3 ones when rolling a dice for 10 times.

Features:

- $m_X(t) = (q + pe^t)^n, \quad m_X : \mathbb{R} \to \mathbb{R}, \quad q = 1 - p$
- $E[X] = np$
- $\text{Var}[X] = npq$

2.3 Geometric Distribution

Meaning: the probability of requiring x trials to obtain the first success.

$$f_X(x) = P[\text{ not succeed until } x\text{th trial }] = (1-p)^{x-1}p$$

Example: tossing a coin for 6 times to get the first head.

Features:

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

2.4 Pascal Distribution

Meaning: the probability that $x \geq r$ trials are needed to obtain r successes.

Let $r \in \mathbb{N}\backslash\{0\}, x \in \{r, r+1, r+2, \ldots\}$, the distribution function f_X given by

$$f_X(x) = \begin{pmatrix} x-1 \\ r-1 \end{pmatrix} p^r (1-p)^{x-r}, \qquad 0 < p < 1$$

is said to follow a Pascal distribution with parameters p and r.

Example: find the probability of using n tossing to get the third head.

Features:

- $m_X : (-\infty, -\ln q) \to \mathbb{R}, \quad m_X(t) = \frac{(pe^t)^r}{(1-qe^t)^r}, \quad q = 1-p$
- $E[X] = r/p$
- $\operatorname{Var} X = rq/p^2$

2.5 Negative Binomial Distribution

Meaning: the probability that x failures are needed to obtain r successes.

Let $r \in \mathbb{N}\backslash\{0\}, x \in \mathbb{N}$, and distribution function f_X is given by

$$f_X(x) = \left(\begin{array}{c} x+r-1 \\ r-1 \end{array} \right) p^r(1-p)^x = \left(\begin{array}{c} -r \\ x \end{array} \right) (-1)^x p^r (1-p)^x, \qquad 0 < p < 1$$

is said to follow a negative binomial distributionwith parameters p and r.

2.6 Hypergeometric Distribution

meaning: The probability of having exactly x red balls after drawing n balls from the urn that has r red balls and $N-r$ black balls.
Assume that

$$r > n \quad \text{and} \quad N-r > n$$

Then we get

$$\begin{aligned} f_X(x) &= P[\text{ exactly } x \text{ red balls out of } n \text{ selected }] \\ &= \frac{(\#\text{ ways to select } x \text{ out of } r \text{ balls }) \cdot (\#\text{ ways to select } n-x \text{ out of } N-r \text{ balls })}{\#\text{ ways to select } n \text{ out of } N \text{ balls}} \\ &= \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} \end{aligned}$$

Example: A production lot of 200 units has 8 defectives. Find the probability of selecting exactly one defective when we choose a random sample of 10 units.

Features:

- $\mathrm{E}[X] = n\frac{r}{N}$
- $\operatorname{Var} X = n\frac{r}{N}\frac{N-r}{N}\frac{N-n}{N-1}$

Relation with Binomial distribution:

- Binomial distribution: a series of i.i.d Bernoulli trials.
- Hypergeometric distribution: a series of identical but not independent Bernoulli trials.
- For Binomial distribution, $\mathrm{E}[X] = np$, $\operatorname{Var} X = npq$, where

  $$p = \frac{r}{N}, \quad q = \frac{N-r}{N}$$

  so only the variance differs by $\frac{N-n}{N-1}$
- the binomial distribution may be used to approximate the hypergeometric distribution if the sampling fraction n/N is small (< 0.05)

**For Assignment3

Exercise 3.4 A Tricky Question involving the Binomial Distribution
A mathematics textbook has 200 pages on which typographical errors in the equations could occur. Suppose there are in fact five errors randomly dispersed among these 200 pages.

i) What is the probability that a random sample of 50 pages will contain at least one error?
(2 Marks)

ii) How large must the random sample be to assure that at least three errors will be found with 90% probability? (You may use a normal approximation to the binomial distribution.)
(3 Marks)

Solution:

- This follows binomial distribution, as the appearance of typos are i.i.d.
- Each typo chooses a page from the book. The page can be the same.

i)
$$P[N \geq 1] = 1 - P[N = 0] = 1 - \left(\frac{150}{200}\right)^5 \approx 0.76$$

ii)
$$P[X \leq y] = \sum_{x=0}^{y} \left(\begin{array}{c} n \\ x \end{array} \right) p^x (1-p)^{n-k} \approx \Phi\left(\frac{y + 1/2 - np}{\sqrt{np(1-p)}}\right)$$

Denote the sample size by x, we have $n = 5, p = \frac{x}{200}$

$$P[N \geq 3] = 1 - P[N/ \leq 2] = 0.1 = \Phi\left(\frac{2+1/2-\frac{x}{40}}{\sqrt{\frac{x}{40}\cdot\left(1-\frac{x}{200}\right)}}\right)$$
$$\Rightarrow \frac{2+1/2-\frac{x}{40}}{\sqrt{\frac{x}{40}\cdot\left(1-\frac{x}{200}\right)}} \approx -1.28155$$
$$\Rightarrow x \geq 150$$

2.7 Poisson Distribution

Meaning: the probability of x occurrences of random events at a constant rate in a continuous environment.

Assumptions:

(i) Independence: If the intervals $T_1, T_2 \subset [0, t]$ do not overlap, then the numbers of arrivals in these intervals are independent of each other.

(ii) Constant rate of arrivals.

Mathematical Assumptions: There exists a number $\lambda > 0$ (arrival rate) such that for any small time interval of size Δt the following postulates are satisfied:

1. The probability that exactly one arrival will occur in an interval of width Δt is $\lambda \cdot \Delta t + o(\Delta t)$
2. The probability that exactly zero arrivals will occur in the interval is $1 - \lambda \cdot \Delta t + o(\Delta t)$
3. The probability that two or more arrivals occur in the interval is $o(\Delta t)$

After calculation, we find the probability x arrivals in the time interval $[0, t]$ is given by

$$p_x(t) = \frac{(\lambda t)^x}{x!} e^{-\lambda t}, \quad x \in \mathbb{N}$$

Or let $k = \lambda t \in \mathbb{R}, x \in \mathbb{N}$. the density function $f_X : \mathbb{N} \to \mathbb{R}$ given by

$$f_X(x) = \frac{k^x e^{-k}}{x!}$$

is said to follow a Poisson distribution with parameter k

Examples: A healthy individual may have an average white blood cell count of as low as $4500/\mathrm{mm}^3$ of blood. To detect a white-cell deficiency, a $0.001\mathrm{mm}^3$ drop of blood is taken and the number X of white blood cells is found ($k = \lambda t = 4500 \times 0.001 = 4.5$).

Features:

- $m_X : \mathbb{R} \to \mathbb{R}, \qquad m_X(t) = e^{k\left(e^t - 1\right)}$
- $E[X] = k$
- $\operatorname{Var} X = k$

** By checking the variance and expectation, we can know whether the case follows a Poisson distribution.

Relation with Binomial Distribution:
For a Binomial distribution, when n is large and p is small, we can approximate the binomial distribution by a Poisson distribution with parameter

$$k = pn$$

We can prove it mathematically by showing

$$\binom{n}{m} p^m (1-p)^{n-m} \quad \xrightarrow[n \cdot p = k]{n \to \infty} \quad \frac{k^m}{m!} e^{-k}$$