2022 SUMMER

업데이터 통계학 스터디

Chapter 2 – Discrete Random Variables

Random Variables

Random Variables

- Whether an experiment yields qualitative or quantitative outcomes, methods of statistical analysis require that we focus on certain numerical aspects of the data such as a sample mean or sample standard deviation.

- The concept of a random variable allows us to pass from the experimental outcomes themselves to a numerical function of the outcomes.

Random Variables

Definition.

A **random variable** is a function from the sample space Ω to the real numbers.

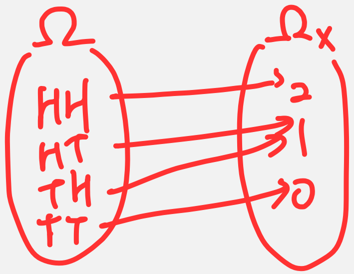
$\Omega \rightarrow \Omega_x$

- A random variable X taking values in a set Ω_X is a function

  $X\colon \Omega \rightarrow \Omega_X$. Ω_X is usually a set of numbers, e.g., $\mathbb{R}$ or $\mathbb{N}$.

- Let $T \subseteq \Omega_X$, define $P(X \in T) = P(\{\omega \in \Omega : \mathrm{X}(\omega) \in T\})$, i.e., the probability that the outcome is in T.

Example 2.1



A random variable is a numerical quantity that is generated by a random experiment. We just assign a numerical number on each possible outcome from the experiment.

Consider to toss a coin twice. Now, we are interested in the number of heads. Let X be the number of heads which is a random number from the random experiment.

$P(X \in A) = P(X=1)$

$= P(\{\omega \in \Omega : X(\omega) = 1\})$

1. The sample space Ω = $\{HH, HT, TH, TT\}$
2. The range of X (all the possible values of X) is $\Omega_X = \{0, 1, 2\}$

Random Variables: Notation

We usually denote random variables by capital letters, such as X, Y, or Z. And the actual values that they can take by lowercase letters, such as x, y, or z.

ex.

Let X: the number of heads in tossing a coin twice.

If we perform the experiment and we observe 2 heads (HH), then $x = 2$.

$\in \Omega_X$

Discrete Random Variables

Discrete Random Variables

$A = \{a \mid 0 \le a \le 10\} \to X$

$B = \{0, 1, 2\} \to$ Discrete

Definition.

- A random variable X is discrete if X can take at most countably many different values. In this case we also say that X has a discrete distribution.

- Each discrete random variable X has a probability mass function (PMF) defined by

확률 질량

함수

$$f_X(x) = P(X = x), \text{ for all real number } x$$

Discrete Random Variables

If X is discrete, there are at most countably many values of x $\in \Omega_X$ such that $f_X(x) > 0$ and the corresponding values of $f_X(x)$ must add to 1.

$$\sum_{i=1}^{\infty} f_X(x_i) = 1$$

Discrete Random Variables: Examples

- Let X be the number of heads when you toss a coin. Then,

  $\Omega = ?$, $\{H, T\}$ $\quad\quad$ $\Omega_X = ?$ $\{0, 1\}$

- Let X be the value shown by rolling a fair die. Then

  $\Omega_X = ?$, $\quad\quad$ PMF? $f_X(x) = P(X = x) = \frac{1}{6}$, $(x \in \Omega_X)$

  $\{1, 2, 3, 4, 5, 6\}$

- Suppose we roll two dice, and let the values obtained by X and Y. Then the sum can be represented by $S = X + Y$, with

  $\Omega_S = ?$ $\{2, \cdots, 12\}$

Cumulative Distribution Function

Every random variable has a Cumulative Distribution function (CDF) defined by

$$F_X(x) = \underline{P(X \le x)}, \text{ for all real number } x,$$

And if X has a discrete distribution with PMF $f_X(x)$,

$$F_X(x) = \sum_{t \le x} f_X(t)$$

Example 2.2

Consider three-coin tosses

$$\{\text{HHH, HHT, HTH, THH, TTH, THT, HTT, TTT}\}.$$

1. Define the random variable X to be the number of heads. $P(X = x)$ for $x = 0, 1, 2, 3$?

X	0	1	2	3
$p(X=x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

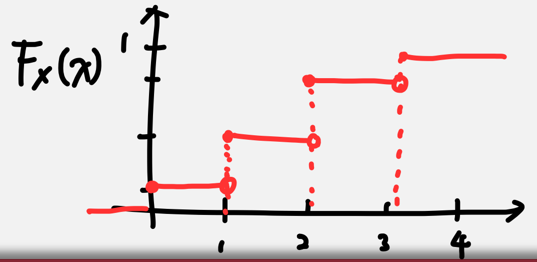
2. Find $F_X(2)$ and $F_X(2.5)$

$F_X(2) = P(X \le 2)$

$= P(X=0) + P(X=1) + P(X=2) = \frac{7}{8}$

$F_X(2.5) = P(X \le 2.5) = \frac{7}{8}$

3. Draw the CDF of X



Cumulative Distribution Function

Every random variable has a CDF that satisfies the following properties:

- $\lim_{x \to -\infty} F_X(x) = 0, \quad \lim_{x \to \infty} F_X(x) = 1, \quad F_X(x)$ is non-decreasing
- $P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F_X(b) - F_X(a)$
- $f_X(x) = P(X = x) = P(X \leq x) - P(X < x)$

Bernoulli trial

Consider a simple experiment with two outcomes: success, with probability p, and failure, with probability $q = (1 - p)$. Such an experiment is called a **Bernoulli trial**.

Let X be a random variable that takes only the values 0 and 1 with $P(X = 1) = p$. The distribution of X is called the Bernoulli distribution with parameter p.

$$X \sim Bernoulli(p)$$

Binomial Distribution

If we perform n independent Bernoulli trials with success probability p, the total number of successes will be Binomial random variable.

A random variable X has a binomial distribution with parameters n and p.

성공 확률

$$X \sim Binomial(\underline{n}, \underline{p})$$

시행 횟수

and it has the PMF

$$f_x(x) = \underline{\binom{n}{x}}\,\underline{p^x}\,\underline{(1-p)^{n-x}}, \text{ for } x = 0, 1, 2, \dots, n$$

Binomial Distribution

The Bernoulli distribution with parameter p, $X \sim Bernoulli(p)$, is the binomial distribution with parameters $n = 1$ and p.

Let X be the number of successes in n independent trials. Then $S_n = \sum_{i=1}^{n} Z_i$ is distributed according to the binomial distribution,

$$X = \sum_{i=1}^{n} Z_i \sim Binomial(n, p), \qquad Z_i \sim Bernoulli(p)$$ (iid)

identically
independent
distribution

Moments and Various Discrete R.V.

정률

Expectation

There are several one-dimensional functionals of a distribution that play important roles in probability. The first of these functionals is the mean or expectation or expected value.

The mean of a discrete random variable X with PMF $f_x(x)$ is

$$E(X) = \sum xP(X = x) = \sum xf_X(x)$$

(If $\sum |x| f_X(x) < \infty$)

Expectation

- If X is a discrete random variable that takes on one of the values x_i, $i \geq 1$, with respective probabilities $f_x(x_i)$, then for any real-valued function g, Y = g(X).

$$E(Y) = E\big(g(X)\big) = \sum g(x_i) f_x(x_i)$$

- If $Y = aX + b$, then $E(Y) = \mathrm{aE(X) + b}$

Example 2.3

If $X \sim Binomial(n, p)$, then find the expectation value of random variable X.

$E(X) = np$

$f_X(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, \cdots, n$

$$E(X) = \sum_{x=0}^{n} x \cdot f_X(x) = \sum_{x=0}^{n} x \cdot \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

$$= \sum_{x=1}^{n} \frac{np\,(n-1)!}{(x-1)!(n-x)!} p^{n-1} (1-p)^{n-x} = np \underbrace{\sum_{y=0}^{n-1} \frac{(n-1)!}{y!(n-1-y)!} p^{n-1} (1-p)^{n-1-y}}_{=1}$$

$$= np$$

Moments

$V(X) = E((X-\mu)^2) = E(X^2) - \{E(X)\}^2$

For each random variable X, and every positive integer k, the expectation $E(X^k)$ is called the *kth moment of X*. Suppose that X is a random variable for which $E(X) = \mu$.

For every integer k, the expectation $E[(X-\mu)^k]$ is called the *kth central moment of X* or *kth moment of X* about the mean.

- The 2nd central moment of X is variance, a measure of how much a distribution is spread out around its mean.

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

Moments

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

The *standard deviation of X* is the sqrt of its variance.

$$\sigma(X) = \sqrt{Var(X)}$$

Moment Generating Function (MGF)

The moment generating function $M_X(t)$ of a discrete random variable X is defined for all real values of t by

$$M_X(t) = E(e^{tX}) = \sum_x e^{tX} f_X(x)$$

Where $f_X(x)$ is the PMF of X

Example 2.4

Find the MGF of discrete random variables X and Y. where,

1) $X \sim Bernoulli(p)$

$M_X(t) = E(e^{tX}) = \sum_{x=0}^{1} e^{tx} \cdot f_X(x) = \sum_{x=0}^{1} e^{tx} \cdot p^x \cdot (1-p)^{1-x} = (1-p)e^0 + pe^t$

$\therefore M_X(t) = pe^t + (1-p)$

2) $Y \sim Binomial(n, p)$

$n=1$

$M_Y(t) = E(e^{tY}) = \sum_{y=0}^{n} e^{ty} \cdot p^y \cdot (1-p)^{n-y} = \sum_{y=0}^{n} (pe^t)^y (1-p)^{n-y} = (pe^t + (1-p))^n$

$\therefore M_Y(t) = (pe^t + p - 1)^n$

Property of MGF

$\frac{d^k}{dt^k} E(e^{tX}) = E\left(\frac{d^k}{dt^k} e^{tX}\right) = E(X^k e^{tX})$

$\frac{d^k}{dt^k} E(e^{tX}) \Big|_{t=0} = E(X^k)$

- If the MGF of X is finite in an open interval around 0, then $E(X^k)$ exists for all $k = 1, 2, 3, \dots$ and

$$E\left(X^k\right) = \frac{d^k}{dt^k} M_X(t) \Big|_{t=0}$$

- If two distribution has the same MGF, then those two $r.v.$ are $i.i.d$ $r.v.$

Example 2.5

$E(X)$ $E(X^2) - \{E(X)\}^2$

Let $X \sim Bernoulli(p)$. Find an expectation value of X and a variance of X, by using MGF property.

$M_X(t) = pe^t + (1-p)$

$Y \sim Bin(n, p)$

$E(Y) = np$, $V(Y) = npq = np(1-p)$

① $E(X) = \frac{d}{dt} M_X(t) \Big|_{t=0} = pe^t \Big|_{t=0} = p$

② $E(X^2) = \frac{d^2}{dt^2} M_X(t) \Big|_{t=0} = pe^t \Big|_{t=0} = p$

$V(X) = p - p^2 = p(1-p)$

Geometric Distribution

1th 2th $\cdots$ $(n-1)$th nth

$1-p$ $1-p$ $1-p$ p $\rightarrow (1-p)^{n-1} \cdot p$

Suppose that you toss a coin until you get a head. Let X be the number of trials until the first head. Suppose that the coin tosses are $i.i.d.$ Bernoulli random variables with parameter p. Find the PMF of X.

A discrete random variable X whose PMF is given by

$$f_X(x) = \begin{cases} (1-p)^{x-1}p & if\ x = 1, 2, \ldots \\ 0 & o.w. \end{cases}$$

otherwise

Is said to have a geometric distribution with parameter p.

$$X \sim Geometric(p)$$

Example 2.6

Let $X \sim Geometric(p)$.

1) Find an expectation value of X and a variance of X.

$\frac{1}{p}$

2) Find a MGF of X.

$M_X(t) = E(e^{tX}) = \sum_{x=1}^{\infty} e^{tx} \cdot (1-p)^{x-1} \cdot p = \sum_{x=1}^{\infty} (e^t(1-p))^x \frac{p}{1-p}, \quad (e^t(1-p) < 1)$

$= \frac{p}{1-p} \cdot \frac{e^t(1-p)}{1-e^t(1-p)} = \frac{pe^t}{1-(1-p)e^t}$

Memoryless property

A discrete distribution has the memoryless property if a random variable X has a distribution satisfying

$$P(X > m + \overset{3}{\overset{\shortparallel}{n}} \mid X > m) = P(X > n)$$

$$\underset{5}{\shortparallel} \qquad \underset{2}{\shortparallel} \qquad \underset{3}{\shortparallel}$$

for all non-negative integers m, n.

$$p(X \le k) = \sum_{x=1}^{k} (1-p)^{k-1} \cdot p$$

$$= \frac{p\{1-(1-p)^k\}}{1-(1-p)} = 1-(1-p)^k$$

$$\frac{p(X > m+n,\ X > m)}{p(X > m)} = \frac{p(X > m+n)}{p(X > m)} = \frac{(1-p)^{m+n}}{(1-p)^m} = (1-p)^n = p(X > n)$$

$$p(X > k) = (1-p)^k$$

Example 2.7

A baseball player's batting average is 0.3.

$* \; p(X>k) = (1-p)^k$

Given that the player has not had a hit after three times at bat, what is the probability the player will not get a hit after five times at bat?

$p(X>5 \mid X>3) = p(X>2)$

x x x _ _

Example 2.8

An urn contains N white and M black balls. Balls are randomly selected, one at a time, until a black one is obtained. If we assume that each ball selected is replaced before the next one is drawn, what is the probability that

$$p = \frac{M}{N+M}$$

1. Exactly n draws are needed? $(1-p)^{n-1} \cdot p$
2. At least k draws are needed?

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

Negative Binomial Distribution

Let X be the number of trials until the rth success occurs with success probability p. Here r is a positive integer greater than or equal to one.

Then, the PMF of X is given by

$$\underbrace{1 \quad 2 \quad \cdots \quad r-1}_{r-1} \qquad n$$

$$f_X(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, \qquad x = r, r+1,$$

When $r = 1$, X has a geometric distribution. For a general r, we say that X has a negative binomial distribution. $X \sim Negbin(r, p)$

Negative Binomial Distribution

Let $Y \sim Negbin(r,p)$

Then we can write $Y = X_1 + \cdots + X_r$ where X_i are $i.i.d.$ as $Geo(p)$.
The expected waiting time to complete r successes is

$$E(X) = \sum_{i=1}^{r} \overset{\frac{1}{p}}{E(X_i)} = \underset{\frac{1}{p} \times r}{\frac{r}{p}}$$

Waiting time

Poisson Distribution

The PMF of the Poisson distribution with parameter λ is

$$X \sim Poisson(\lambda), \qquad f_X(x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!} \qquad for\ x = 0, 1, 2, \dots$$

The parameter λ must be positive.

Example 2.9

Let $X \sim Poisson(\lambda)$. Find that

1. Expected value of X

$M_x'(t) = \lambda e^t \cdot e^{\lambda(e^t-1)}$, $E(X) = M_x'(0) = \lambda$

2. Variance of X

$M_x''(t) = \lambda e^t \cdot e^{\lambda(e^t-1)} + \lambda^2 e^{2t} \cdot e^{\lambda(e^t-1)}$, $E(X^2) = M_x''(0) = \lambda^2 + \lambda$

$Var(X) = E(X^2) - \{E(X)\}^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$

3. MGF of X

$$M_x(t) = E(e^{tX}) = \sum_{x=0}^{\infty} e^{tx} \cdot \frac{e^{-\lambda} \cdot \lambda^x}{x!}$$

$$= e^{-\lambda} \cdot \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!} = e^{-\lambda} \cdot e^{\lambda e^t} = e^{\lambda(e^t-1)}$$

Formal Definition of Poisson Dist.

Poisson distribution arises in situations where "events" occur at certain points in time. Poisson event is the one that satisfies three following assumptions:

(i) The probability that exactly one event occurs in a given interval of length h is equal to $\lambda h + o(h)$;

(ii) The probability that more than two events occur in an interval of length h is equal to $o(h)$;

(iii) The occurrences of the events for any non-overlapping intervals are independent,

Formal Definition of Poisson Dist.

where $o(h)$ (called small o) stands for any function $f(h)$ for which $\lim_{h\to 0} f(h)/h = 0$

If we let $N(t)$ be the number of the Poisson events occurring in the interval $[0, t]$, then

$$N(t) \sim Poisson(\lambda t)$$

Poisson approximation to Binomial

$X \sim Bin(n, p)$, If $n \uparrow$ and $p \downarrow$, $\lambda = np \rightarrow X \sim Poisson(\lambda)$

Example 2.10

In a manufacturing process where glass products are made, defects or bubbles occur, occasionally rendering the piece undesirable for marketing. It is known that, on average, 1 in every 1000 of these items produced has one or more bubbles.

$p = \frac{1}{1000}$

What is the probability that a random sample of 8000 will yield fewer than 7 items possessing bubbles?

$n = 8000$

$X \sim Bin\left(8000, \frac{1}{1000}\right) \sim Poisson(8)$

$f_X(x) = \frac{e^{-8} \cdot 8^x}{x!}$

$P(X \le 6) = \sum_{x=0}^{6} \frac{e^{-8} \cdot 8^x}{x!} \approx 0.3134$