Random Variable (r. v.) Chap. 3

A r. v. X is a function that assigns a real number, $X(\xi)$ to each outcome ξ in the sample space of a random experiment. The sample space S is the domain of the random variable and the set S_X of all values taken on by X is the range of the r.v.. Note S_X is a subset of of the set of all real numbers.

EX.3.1 e.g. A coin is tossed 3 times and the sequence of heads and tails is noted. The sample space is

$$\{HHH, HHT, HTH, \cdots, TTT\}$$

- Let X be a r.v. which the number of heads in 3 tosses

so ξ: HHH HHT HTH H...

$X(\xi)$: 3 2 2

---

e.g. 2 A pair of fair dice is tossed, the sample space is

$S = \{(1,1),(1,2), \cdots (6,6)\}$

Let X be a r.v. that $X(a,b) = \max(a,b)$

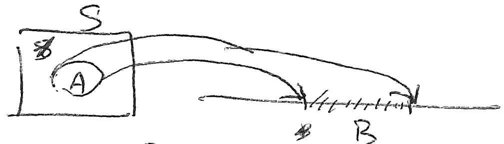
then $S_X = \{1, 2, 3, 4, 5, 6\}$

A discrete r.v. X is defined as a r.v. that assumes values from a countable set, i.e., $S_X = \{x_1, x_2, \ldots \ldots\}$

It is finite if $S_X = \{x_1, x_2, \ldots, x_n\}$

Probability Mass function PMF of a discrete r.v. X is defined as

$$p_X(x) = P[X = x] = P[\{\zeta : X(\zeta) = x\}] \quad \text{for } x \text{ a real number}$$



Properties of PMF

$$P[X \in B] = P[A] = P[\{\zeta : X(\zeta) \in B\}]$$

1) $p_X(x) \geq 0$ for all x

2) $$\sum_{x \in S_X} p_X(x) = \sum_{\text{all } k} p_X(x_k) = \sum_{\text{all } k} P[A_k] = 1$$

3) $$P[X \text{ in } B] = \sum_{x \in B} p_X(x) \quad \text{where } B \subset S_X$$

Example 3.1 Coin Tosses

A coin is tossed three times and the sequence of heads and tails is noted. The sample space for this experiment is $S = \{\text{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}\}$. Let X be the number of heads in the three tosses. X assigns each outcome ζ in S a number from the set $S_X = \{0, 1, 2, 3\}$. The table below lists the eight outcomes of S and the corresponding values of X.

ζ:	HHH	HHT	HTH	THH	HTT	THT	TTH	TTT
$X(\zeta)$:	3	2	2	2	1	1	1	0

X is then a random variable taking on values in the set $S_X = \{0, 1, 2, 3\}$.

Example 3.2 A Betting Game

A player pays \$1.50 to play the following game: A coin is tossed three times and the number of heads X is counted. The player receives \$1 if $X = 2$ and \$8 if $X = 3$, but nothing otherwise. Let Y be the reward to the player. Y is a function of the random variable X and its outcomes can be related back to the sample space of the underlying random experiment as follows:

ζ:	HHH	HHT	HTH	THH	HTT	THT	TTH	TTT
$X(\zeta)$:	3	2	2	2	1	1	1	0
$Y(\zeta)$:	8	1	1	1	0	0	0	0

Y is then a random variable taking on values in the set $S_Y = \{0, 1, 8\}$.

Example 3.3 Coin Tosses and Betting

Let X be the number of heads in three independent tosses of a fair coin. Find the probability of the event $\{X = 2\}$. Find the probability that the player in Example 3.2 wins \$8.

Note that $X(\zeta) = 2$ if and only if ζ is in $\{\text{HHT, HTH, THH}\}$. Therefore

$$\begin{aligned} P[X = 2] &= P[\{\text{HHT, HTH, HHT}\}] \\ &= P[\{\text{HHT}\}] + P[\{\text{HTH}\}] + P[\{\text{HHT}\}] \\ &= 3/8. \end{aligned}$$

The event $\{Y = 8\}$ occurs if and only if the outcome ζ is HHH, therefore

$$P[Y = 8] = P[\{\text{HHH}\}] = 1/8.$$

Some important random variables

) Bernouli Random Variable

Let S be the sample space of a random experiment and A be an event of this experiment. The r.v. I defined as:

$$I(\xi) = \begin{cases} 0 & \text{if } \xi \text{ not in } A \\ 1 & \text{if } \xi \text{ is in } A \end{cases}$$

is called Bernouli r.v. with range $S_X = \{0, 1\}$ and its pmf $p(0) = 1 - p$
$p(1) = p$

Binomial R. V.

2) suppose a random experiment is repeated n times. Let X be the no. of times a certain event A occurs in these n trials (e.g. X could be the no. of heads in n tosses of a coin)

$$p(X=k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \ldots, n$$

$$S_X = \{0, 1, \ldots, n\}$$

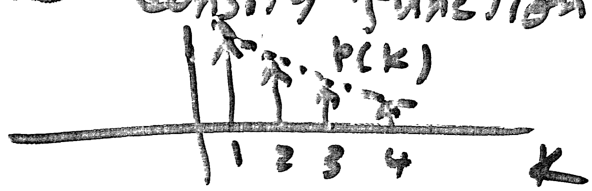
3) Geometric R. V.

The no. M of Bernoulli trials until the first occurrence of a success is a r. v. Called Geometric r. v.

$$p(M=k) = (1-p)^{k-1} p, \quad k = 1, 2, \ldots$$

Where p is the prob. of success

The density function decays geometrically



$$P(M \le k) = \sum_{j=1}^{k} p q^{j-1} = p \frac{1-q^k}{1-q} = 1 - q^k$$

Where $q = 1 - p$

Poisson R. V.

Poisson r.v. arises in situation where the events occur in time or space (e.g. the no. of auto accidents at a particular intersection during a time period of one week).

Subdivide 1 week period into n subinterval, each of which is so small that at ~~least~~ most one accident could occur in it with nonzero prob. p. Note, as n increase, p, the prob. of one accident of these shorter intervals will decrease. Let $\lambda = np$ and taking the limit of binomial prob.

$$p(y) = \binom{n}{y} p^y (1-p)^{n-y} \quad \text{as } n \to \infty$$

$$p(y) = \lim_{n\to\infty} \binom{n}{y} p^y (1-p)^{n-y} = \lim_{n\to\infty} \frac{n(n-1)(n-2)\cdots(n-y+1)}{y!} \cdot \left(\frac{\lambda}{n}\right)^y \left(1-\frac{\lambda}{n}\right)^{n-y}$$

$$= \lim \frac{\lambda^y}{y!}\left(1-\frac{\lambda}{n}\right)^n \frac{\cancel{n}(n-1)\cdots(n-y+1)}{n^{y-1}} \left(1-\frac{\lambda}{n}\right)^{-y}$$

$$= \frac{\lambda^y}{y!} \lim_{n\to\infty} \underbrace{\left(1-\frac{\lambda}{n}\right)^n}_{e^{-\lambda}} \underbrace{\left(1-\frac{\lambda}{n}\right)^{-y}}_{=1} \underbrace{\left(1-\frac{1}{n}\right)}_{=1} \underbrace{\left(1-\frac{2}{n}\right)}_{=1} \cdots \underbrace{\left(1-\frac{y-1}{n}\right)}_{=1}$$

$$p(y) = \frac{\lambda^y}{y!} e^{-\lambda}, \quad y = 0, 1, 2, \ldots$$

Note:

$$e^{x} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

$$= \sum_{y=0}^{\infty} \frac{x^y}{y!}$$

so, $$\sum_{y=0}^{\infty} p(y) = \sum_{y=0}^{\infty} \frac{\lambda^y}{y!} e^{-\lambda} = e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^y}{y!}$$

$$= e^{-\lambda} \cdot e^{\lambda} = 1$$

Our text use α instead of λ

so, $$\boxed{p_k = \frac{\alpha^k}{k!} e^{-\alpha}, \quad k = 0, 1, 2, \ldots \qquad \alpha = np}$$

Expected value of and moments of Discrete r.v.

$$m_X = E[X] = \sum_{x \in S_X} x P_X(x) = \sum_k x_k P_X(x_k)$$

Mean of a Bernoulli r.v. (I_A)

$$E[X] = 0\, P_I(0) + 1\, P_I(1) = 0(1-p) + 1(p) = p$$

where p is the probability of success in the Bernoulli trial

Mean of a uniform discrete r.v.

$$E[X] = \sum_{k=0}^{M-1} k \frac{1}{M} = \frac{1}{M}\left[0+1+2+\cdots M-1\right] = \frac{M(M-1)}{2M}$$

$$= \frac{M-1}{2}$$

Expected value of Functions of a r.v.

$$E[g(X)] = \sum_k g(x_k) P_X(x_k)$$

Variance of a r.v.

$$\sigma_X^2 = VAR[X] = E[(X-m_X)^2]$$

$$= \sum_{x \in S_X} (x - m_X) P_X(x) = \sum_{k=1}^{\infty} (x_k - m_X)^2 P_X(x_k)$$

$$VAR[X] = E[(X-m_X)^2] = E[X^2 - 2m_X X + m_X^2] = E[X^2] - m_X^2$$

$$VAR[X+c] = E[(X+c-(E[X]+c))^2]$$
$$= E[(X-E[X])^2] = VAR[X]$$

$$VAR[cX] = E[(cX - cE[X])^2]$$
$$= E[c^2(X-E[X])^2]$$
$$= c^2 VAR[X]$$

variance of Bernoulli r. v.

$$VAR[I_A] = E[I_A^2] - (E[I_A])^2$$

$$E[I_A^2] = 0^2(1-p) + 1^2(p) = p$$

so $VAR[I_A] = p - p^2 = p(1-p) = pq$

Examples 3-12, 3-20

Mean of poisson r. v.

$$E[X] = \sum_{k=0}^{\infty} k P(X=k)$$
$$= \sum_{k=1}^{\infty} k \frac{\lambda^k e^{-\lambda}}{k!}$$
$$= \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} e^{\lambda} = \lambda$$

Mean of Geometric r.v.

$$E[X] = \sum_{k=0}^{\infty} k(1-p)p^k$$
$$= p(1-p)\sum_{k=0}^{\infty} \frac{d(p^k)}{dp}$$
$$= p(1-p)\frac{d}{dp}\frac{1}{1-p}$$
$$= \frac{p}{1-p}$$

Conditional probability mass function

$$p_X(x \mid C) = P[X = x \mid C] = \frac{P[\{X = x\} \cap C]}{P[C]}$$

Conditional Expected value

$$m_{X|B} = E[X \mid B] = \sum_{x \in S_X} x \, p_X(x \mid B) = \sum_k x_k p_X(x_k \mid B)$$

Conditional Variance of X given B

$$VAR[X \mid B] = E[(X - m_{X|B})^2 \mid B]$$

$$= \sum_{k=1}^{\infty} [(x_k - m_{X|B})^2 p_X(x_k \mid B)$$

$$= E[X^2 \mid B] - m_{X|B}^2$$

Example 3.5 Coin Tosses and Binomial Random Variable

Let X be the number of heads in three independent tosses of a coin. Find the pmf of X.

Proceeding as in Example 3.3, we find:

$$p_0 = P[X = 0] = P[\{TTT\}] = (1 - p)^3,$$
$$p_1 = P[X = 1] = P[\{HTT\}] + P[\{THT\}] + P[\{TTH\}] = 3(1 - p)^2 p,$$
$$p_2 = P[X = 2] = P[\{HHT\}] + P[\{HTH\}] + P[\{THH\}] = 3(1 - p)p^2,$$
$$p_3 = P[X = 3] = P[\{HHH\}] = p^3.$$

Note that $p_X(0) + p_X(1) + p_X(2) + p_X(3) = 1.$

Example 3.6 A Betting Game

A player receives \$1 if the number of heads in three coin tosses is 2, \$8 if the number is 3, but nothing otherwise. Find the pmf of the reward Y.

$$p_Y(0) = P[\zeta \in \{TTT, TTH, THT, HTT\}] = 4/8 = 1/2$$
$$p_Y(1) = P[\zeta \in \{THH, HTH, HHT\}] = 3/8$$
$$p_Y(8) = P[\zeta \in \{HHH\}] = 1/8.$$

Note that $p_Y(0) + p_Y(1) + p_Y(8) = 1.$

Example 3.11 Mean of Bernoulli Random Variable

Find the expected value of the Bernoulli random variable I_A.

From Example 3.8, we have

$$E[I_A] = 0p_I(0) + 1p_I(1) = p.$$

where p is the probability of success in the Bernoulli trial.

Example 3.12 Three Coin Tosses and Binomial Random Variable

Let X be the number of heads in three tosses of a fair coin. Find $E[X]$.

Equation (3.8) and the pmf of X that was found in Example 3.5 gives:

$$E[X] = \sum_{k=0}^{3} k p_X(k) = 0\left(\frac{1}{8}\right) + 1\left(\frac{3}{8}\right) + 2\left(\frac{3}{8}\right) + 3\left(\frac{1}{8}\right) = 1.5.$$

Note that the above is the $n = 3$, $p = 1/2$ case of a binomial random variable, which we will see has $E[X] = np$.

Example 3.15 Mean of a Geometric Random Variable

Let X be the number of bytes in a message, and suppose that X has a geometric distribution with parameter p. Find the mean of X.

X can take on arbitrarily large values since $S_X = \{1, 2, \dots\}$. The expected value is:

$$E[X] = \sum_{k=1}^{\infty} kpq^{k-1} = p\sum_{k=1}^{\infty} kq^{k-1}.$$

This expression is readily evaluated by differentiating the series

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k \tag{3.13}$$

to obtain

$$\frac{1}{(1-x)^2} = \sum_{k=0}^{\infty} kx^{k-1}. \tag{3.14}$$

Letting $x = q$, we obtain

$$E[X] = p\frac{1}{(1-q)^2} = \frac{1}{p}. \tag{3.15}$$

We see that X has a finite expected value as long as $p > 0$.

Example 3.22 Variance of Geometric Random Variable

Find the variance of the geometric random variable.

Differentiate the term $(1 - x^2)^{-1}$ in Eq. (3.14) to obtain

$$\frac{2}{(1-x)^3} = \sum_{k=0}^{\infty} k(k-1)x^{k-2}.$$

Let $x = q$ and multiply both sides by pq to obtain:

$$\frac{2pq}{(1-q)^3} = pq\sum_{k=0}^{\infty} k(k-1)q^{k-2}$$

$$= \sum_{k=0}^{\infty} k(k-1)pq^{k-1} = E[X^2] - E[X].$$

So the second moment is

$$E[X^2] = \frac{2pq}{(1-q)^3} + E[X] = \frac{2q}{p^2} + \frac{1}{p} = \frac{1+q}{p^2}$$

$$\text{VAR}[X] = E[X^2] - E[X]^2 = \frac{1+q}{p^2} - \frac{1}{p^2} = \frac{q}{p^2}.$$

/

Example 3.20 Three Coin Tosses

Let X be the number of heads in three tosses of a fair coin. Find VAR[X].

$$E[X^2] = 0\left(\frac{1}{8}\right) + 1^2\left(\frac{3}{8}\right) + 2^2\left(\frac{3}{8}\right) + 3^2\left(\frac{1}{8}\right) = 3 \quad \text{and}$$
$$\text{VAR}[X] = E[X^2] - m_X^2 = 3 - 1.5^2 = 0.75.$$

Recall that this is an $n = 3$, $p = 1/2$ binomial random variable. We see later that variance for tl binomial random variable is npq.

Example 3.28 Variance of a Binomial Random Variable

To find $E[X^2]$ below, we remove the $k = 0$ term and then let $k' = k - 1$:

$$\begin{aligned}
E[X^2] &= \sum_{k=0}^{n} k^2 \frac{n!}{k!(n-k)!} p^k(1-p)^{n-k} = \sum_{k=1}^{n} k \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k} \\
&= np \sum_{k'=0}^{n-1} (k'+1) \binom{n-1}{k'} p^{k'}(1-p)^{n-1-k'} \\
&= np\left\{ \sum_{k'=0}^{n-1} k' \binom{n-1}{k'} p^{k'}(1-p)^{n-1-k'} + \sum_{k'=0}^{n-1} 1 \binom{n-1}{k'} p^{k'}(1-p)^{n-1-k'} \right\} \\
&= np\{(n-1)p + 1\} = np(np + q).
\end{aligned}$$

$k' = k - 1$

$k = k' + 1$

In the third line we see that the first sum is the mean of a binomial random variable with parameters $(n - 1)$ and p, and hence equal to $(n - 1)p$. The second sum is the sum of the binomial probabilities and hence equal to 1.

We obtain the variance as follows:

$$\sigma_X^2 = E[X^2] - E[X]^2 = np(np + q) - (np)^2 = npq = np(1 - p).$$

We see that the variance of the binomial is n times the variance of a Bernoulli random variable. We observe that values of p close to 0 or to 1 imply smaller variance, and that the maximum variability is when $p = 1/2$.

Example 3.27 Mean of a Binomial Random Variable

The expected value of X is:

$$\begin{aligned}
E[X] &= \sum_{k=0}^{n} k p_X(k) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=1}^{n} k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\
&= np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-k)!} p^{k-1}(1-p)^{n-k} \\
&= np \underbrace{\sum_{j=0}^{n-1} \frac{(n-1)!}{j!(n-1-j)!} p^j (1-p)^{n-1-j}}_{=1} = np,
\end{aligned} \tag{3.35}$$

$j = k - 1$

where the first line uses the fact that the $k = 0$ term in the sum is zero, the second line cancels out the k and factors np outside the summation, and the last line uses the fact that the summation is equal to one since it adds all the terms in a binomial pmf with parameters $n - 1$ and p.