Math 3280 20-09-24

Review.

Independence of 2 or more events

- Say two events E and F are independent if

$$P(EF) = P(E) \cdot P(F).$$

- $E_1, E_2, \cdots, E_n$ are said to be independent if

$$P(E_{i_1} \cdots E_{i_r}) = P(E_{i_1}) \cdots P(E_{i_r})$$

for all subsets of Events $E_{i_1}, \cdots, E_{i_r}$.

- A family of events $\{E_i\}_{i \in \mathcal{I}}$ are said to be independent if

$\boxed{\text{any finite subfamily}}$ of these events are independent.

Def. (Independence of sub-experiments)

An experiment might consist of some sub-experiments.

For instance, the experiment that rolling a coining continuously consists of a sequence of sub-experiments, where the n-th subexperiment is the n-th toll of the coin, $n = 1, 2, \cdots$.

We say these sub-experiments are independent if

E_1 $E_2, \cdots, E_n$ are independent

whenever E_i is an event whose occurence depends only on the i-th sub-experiment.

These sub-experiments are said to be trials if the set of possible outcomes of each sub-experiment are the same.

Chap 4. Random variables.

§4.1 Introduction to random variables.

Def. For a random experiment, a random variable (r.v.) X is a real-valued function defined on the sample space S. That is,

$$X : S \longrightarrow \mathbb{R}.$$

is a function.

Example 1. Flip 3 fair coins. Let X be the number of the heads that appear.

$$X = \#\{\text{heads that appear}\}.$$

Eg. if the outcome is (T, H, T), then $X = 1$

if the outcome is (H, T, H), then $X = 2$.

Example 2. Two fair dice are rolled.

Y = the product of the numbers that appear.

If the outcome is $(2,5)$, then $Y=10$.

Example 3.

Z = the life-time (in hours) of a randomly chosen flashlight.

§ 4.2. Discrete random Variables.

Def. A r.v. X is said to be discrete if X can take on at most countably many different values.

Def. For a discrete r.v. X, the prob. mass function of X is defined by

$$p(a) = P\{X = a\}$$
$$= P\{\omega \in S : X(\omega) = a\},$$
$$\forall\, a \in \mathbb{R}.$$

Example 4: $X = \#\{$ heads that appear in rolling 3 fair coins $\}$

$$\{X=0\} = \{(T,T,T)\}$$
$$\{X=1\} = \{(H,T,T), (T,H,T), (T,T,H)\}$$
$$\{X=2\} = \{(H,H,T), (H,T,H), (T,H,H)\}.$$
$$\{X=3\} = \{(H,H,H)\}.$$

So $\quad p(0) = \frac{1}{8}, \quad p(1) = \frac{3}{8}, \quad p(2) = \frac{3}{8}, \quad p(3) = \frac{1}{8}$

and $p(a) = 0$ for all $a \notin \{0, 1, 2, 3\}$.

Prop. Let $x_1, x_2, \cdots$ be the possible values of a discrete r.v. X. Then

$$\sum_n p(x_n) = 1.$$

In general $\sum_{x:\, p(x)>0} p(x) = 1$.

Pf. Let

$$E_n = \{ \omega \in S : X(\omega) = x_n \}, \quad n = 1, 2, \cdots$$

Then $E_1, E_2, \cdots$, are mutually exclusive and exhaustive.

Hence $\sum_n P(E_n) = P(S) = 1$.

But $p(x_n) = P(E_n)$, hence

$$\sum_n p(x_n) = 1. \quad \square$$

§ 4.3 Expected value.

Let X be a discrete r.v.

Let $p(x) = P\{X = x\}$ be the prob. mass function of X.

Def. The expected value of X is defined by

$$E[X] = \sum_{x:\, p(x)>0} x \cdot p(x)$$

We sometimes also call $E[X]$ the mean of X.

Hence $E[X]$ is a weighted average of the possible values of X.

Example: $X = \#\{$heads that appear in flipping 3 fair coins$\}$

$p(0) = 1/8$, $p(1) = 3/8$, $p(2) = 3/8$, $p(3) = \frac{1}{8}$

Hence by definition,

$$E[X] = 0\times p(0) + 1\times p(1) + 2\times p(2) + 3\times p(3)$$

$$= 0\times\frac{1}{8} + 1\times\frac{3}{8} + 2\times\frac{3}{8} + 3\times\frac{1}{8}$$

$$= \frac{3+6+3}{8} = 3/2 .$$

§4.4. Expected value of a function of a r.v.

Let $X: S \to \mathbb{R}$ be a discrete r.v.

Let $g: \mathbb{R} \to \mathbb{R}$ be a function.

Then $g(X)$ is a function from S to $\mathbb{R}$, so it is a new r.v.

Q: How can we compute $E[g(X)]$?

Remark: By def, if $y_1, y_2, \cdots,$ are the possible values of $g(X)$, then

$$E[g(X)] = \sum_j y_j P\{g(X) = y_j\}.$$

Prop. $E[g(X)] = \sum_i g(x_i) \cdot p(x_i)$,

where $x_1, x_2, \cdots,$ are all the possible values of X.

Pf. Grouping all $g(x_i)$ which take the same value, gives

$$\sum_i g(x_i) p(x_i)$$

$$= \sum_j \left(\sum_{i:\, g(x_i) = y_j} g(x_i) p(x_i) \right)$$

$$= \sum_j \left(\sum_{i:\, g(x_i) = y_j} y_j p(x_i) \right)$$

$$= \sum_j y_j \left(\sum_{i:\, g(x_i) = y_j} p(x_i) \right)$$

$$= \sum_j y_j P(g(X) = y_j)$$

This proves our desired identity. $\blacksquare$

$$\{ g(X) = y_j \} = \bigcup_{i : g(x_i) = y_j} \{ X = x_i \}$$

(the union being disjoint)

Hence
$$P\{ g(X) = y_j \} = \sum_{i : g(x_i) = y_j} P\{ X = x_i \}$$
$$= \sum_{i : g(x_i) = y_j} p(x_i).$$

Corollary: $E[aX+b] = a\,E[X] + b$

$\forall\, a, b \in \mathbb{R}$

and X is a discrete r.v.

Pf. Let $g : \mathbb{R} \to \mathbb{R}$ be defined by $g(x) = ax+b$.

$$\text{LHS} = E[g(X)] = \sum_{x : p(x) > 0} (ax+b)\, p(x)$$

$$= \sum_{x : p(x) > 0} ax\,p(x) + b\,p(x)$$

$$= a \sum_{x : p(x) > 0} x\,p(x) + b \sum_{x : p(x) > 0} p(x)$$

$$= a \cdot E[X] + b = \text{RHS}. \quad \square$$

§ 4.5 Variance.

Def. Let X be a discrete r.v.

Set

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

we call $Var(X)$ the variance of X

It describes how far X is spread out from its mean.