RANDOM VARIABLES

Random variable = function of an outcome

$$X = f(outcome)$$

$$S \rightarrow \begin{cases} \text{real numbers} \\ \text{integers} \\ (0,1) \\ (0,+\infty) \\ \text{etc.} \end{cases}$$

(domain $\rightarrow$ range)

Random variable is a quantity that depends on chance.

A value of a random variable becomes known once an experiment is completed and its outcome is obtained.

Example: toss 3 fair coins.

Let X = number of heads

Possible values: $\{0, 1, 2, 3\}$, and

$$\begin{aligned} P\{X = 0\} &= 1/8 \\ P\{X = 1\} &= 3/8 \\ P\{X = 2\} &= 3/8 \\ P\{X = 3\} &= 1/8 \end{aligned}$$

(Same model for good/defective items, pass/fail, girl/boy = *Bernoulli trials*)

Summarize:

x	Probability mass function pmf $p_X(x) = P\{X = x\}$
0	1/8
1	3/8
2	3/8
3	1/8
$TOTAL$	1

Properties:

$$0 \leq p(x) \leq 1$$

$$\sum_x p(x) = 1$$

RANDOM VARIABLES

(mixed)

DISCRETE	CONTINUOUS
Finite or countable number of values	The whole interval of possible values
Examples	
number of jobs in a queue	execution time
	waiting time
number of errors	temperature
number of failures $(0, 1, 2, ...)$	height, weight
	intensity
proportion of failures $(0, \frac{1}{N}, \frac{2}{N}, ...)$	distance
	miles per gallon
high jump	long jump

Distribution of a random variable X
= collection of probabilities

$$P\{X \in A\} = \sum_{x \in A} p_X(x)$$

Examples: $P\{X = 3\}$, $\{X > 10\}$,
$P\{X$ is an even number $\}$

Cumulative distribution function (CDF) of X is

$$F_X(x) = P\{X \leq x\} = \sum_{y \leq x} p_X(y)$$

Properties:
$F(x)$ is nondecreasing
Jumps by $p(x)$ at the point x
$F(-\infty) = 0, \quad F(+\infty) = 1$

Computing probabilities:

$$P\{a < X \leq b\} = F(b) - F(a)$$

Random vectors and joint distribution

If X, Y = random variables, then

(X, Y) = random vector

It has a **joint pmf**

$$\begin{aligned} p(x, y) &= P\{(X, Y) = (x, y)\} \\ &= P\{X = x \cap Y = y\} \end{aligned}$$

From $p(x, y)$, the marginal pmf of X and Y are

$$\begin{aligned} p_X(x) &= P\{X = x\} = \sum_y p(x, y) \\ p_Y(y) &= P\{Y = y\} = \sum_x p(x, y) \end{aligned}$$

From $p_X(x)$ and $p_Y(y)$, in general, one cannot compute $p(x, y)$.

Independence

Random variables X and Y are **independent** if

$$p(x, y) = p_X(x)p_Y(y)$$

i.e., $\{X = x\}$ and $\{Y = y\}$ are independent events for **all** x and y.

So,

- to show independence, verify this equality for all x and y;

- to show dependence, find one pair (x, y) violating it

EXPECTATION AND MOMENTS

Expectation of a random variable

Let X = random variable.

$\mathbf{E}(X)$ = its *expectation*, the average value, the mean

---

X is random. It takes different values with probabilities $P(x)$.

$\mathrm{E}(X)$ is constant, non-random.

Example 1: Bernoulli(p), $p = 1/2$.

$$X = \begin{cases} 0 & \text{with probability } 1/2 \\ 1 & \text{with probability } 1/2 \end{cases} \Rightarrow \mathbf{E}(X) = 1/2$$

Example 2: Bernoulli(p), $p = 1/3$.

$$X = \begin{cases} 0 & \text{with probability } 2/3 \\ 1 & \text{with probability } 1/3 \end{cases} \Rightarrow \mathbf{E}(X) = 1/3$$

Definition (Discrete distribution):

$$\mu = \mathbf{E}(X) = \sum_x xP(x)$$ (center of mass)

Expectation of a function $Y = g(X)$

$$\mathbf{E}g(X) = \sum_x g(x)P(x)$$

Variance of a random variable

Example. Consider two financial deals.

$$\begin{array}{llllll} \#1. & P(480) & = & P(520) & = & 0.5 \\ \#2. & P(0) & = & P(1000) & = & 0.5 \end{array}$$

Same $\mathbf{E}(X) = 500$.

In #1, values of X are close to $\mathbf{E}(X)$.
Low variability.

In #2, values of X are far from $\mathbf{E}(X)$.
High variability.

Market term: high volatility

Definition

Variance of X = $\text{Var}(X) = \mathbf{E}\{X - \mathbf{E}(X)\}^2$

Discrete case: $\text{Var}(X) = \sum_x (x-\mu)^2 P(x)$

Standard deviation $\sigma = \sqrt{\text{Var}(X)}$.

X, $\mu = \mathbf{E}(X)$, σ are measured in units
$\sigma^2 = \text{Var}(X)$ is measured in *squared units*

Variance of the profit = 1 mln. squared dollars
Variance of the enrollment = 1000 squared students

Properties

$\mathbf{E}(aX + b) = a\,\mathbf{E}(X) + b$ - always

$\mathbf{E}(X + Y) = \mathbf{E}(X) + \mathbf{E}(Y)$ - always

$\mathbf{E}(XY) = \mathbf{E}(X)\,\mathbf{E}(Y)$ - for independent X, Y

$\text{Var}(aX + b) = a^2\,\text{Var}(X)$ - always

$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$
- for independent X, Y

In general,
$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\,\text{Cov}(X, Y)$
- always

Covariance of X and Y

$$\text{Cov}(X,Y) = \mathbf{E}\{X - \mathbf{E}(X)\}\{Y - \mathbf{E}(Y)\}$$

Properties:

$\text{Cov}(X,X) = \text{Var}(X)$

$\text{Cov}(X,Y) = 0$ for independent X, Y

Independent $\Rightarrow$ uncorrelated, but
Uncorrelated $\not\Rightarrow$ independent, in general