Discrete Random Variables 1

STAT-S520

Arturo Valdivia

01-26-23

These slides complement ISI Chapter 4

Discrete random variable

- Let X be a random variable. X is discrete if $\overbrace{X(S)}^{\text{Range of } X}$ is countable
  - A set is countable if it is finite or denumerable
  - A set is denumerable if there is a one-to-one correspondence with $\mathbb{N}$

    set of natural numbers

Exercise 1

- Let $X(S) = \{0, 1, 2\}$, then X is ... discrete
- Let $Y(S) = \mathbb{N}$ (naturals), then Y is ... discrete
- Let $X_1(S) = \mathbb{R}$ (reals), then X_1 is ... Continuous.
- Let $X_2(S) = \mathbb{Z}$ (integers), then X_2 is ... discrete
- Let $X_3(S) = \mathbb{Q}$ (rationals), then X_3 is ... discrete
- Let $Z(S) = [0, 1]$, then Z is ... continuous

Probability Mass Function

Let X be a discrete random variable. The probability mass function (PMF) of X is a function f, with $f : \mathbb{R} \to [0,1]$, such that, for any $y \in \mathbb{R}$:

$$f(y) = \begin{cases} P(\overbrace{X = y}^{\{s \in S : X(s) = y\}}) & \text{if } y \in X(S) \\ 0 & \text{otherwise} \end{cases}$$

Exercise 2

$X(S) = \{0, 1, 2\}$

We toss a fair coin twice and X is the number of heads. What is

- $f(1) =$ $P(X=1) = P(\{s \in S : X(s) = 1\}) = P(\{HT, TH\}) = 0.5$
- $f(2) =$ 0.25
- $f(\pi) =$ 0
- $f(10) =$ 0

CDF and PMF

If X is a discrete random variable with range $X(S)$ and for any given $y \in \mathbb{R}$ the relationship between the PMF and CDF is given by:

$$F(y) = \sum_{x : x \in L(y)} f(x)$$

where $L(y)$ is the set of numbers in the range of X that are less than or equal to y, or in math notation: $L(y) = X(S) \cap (-\infty, y]$,

Exercise 3

We toss a fair coin twice and X is the number of heads. Let's find $F(\sqrt(2))$ using the PMF

$$X(S) = \{0, 1, 2\}$$

$$L(\sqrt{2}) = \{0,1,2\} \cap (-\infty, \sqrt{2}] = \{0, 1\}$$

$$F(\sqrt{2}) = \sum_{x : x \in L(\sqrt{2})} f(x)$$

$$= \sum_{x : x \in \{0,1\}} f(x)$$

$$= f(0) + f(1)$$

$$= 0.25 + 0.5 = 0.75$$

.

Bernoulli trials

- A Bernoulli trial is a random variable where $X(S) = \{0, 1\}$.
  - It is customary to call $X = 1$ a success and $X = 0$ a failure.
- There is a family of Bernoulli trials with parameter p where $p = P(X = 1)$ and we write

$$X \sim \text{Bernoulli}(p)$$

PMF of a Bernoulli trial

- If $X \sim \text{Bernoulli}(p)$, the PMF of X is given by

$$f(x) = \begin{cases} p & \text{if } x = 1 \\ 1 - p & x = 0 \\ 0 & \text{otherwise} \end{cases}$$

- We can simplify this notation by writing

$$\boxed{f(x) = p^x (1-p)^{1-x}}$$

for $x \in \{0, 1\}$ and $f(x) = 0$ otherwise.

Binomial distribution

X_1, X_2 are independent if for any $x_1, x_2 \in \mathbb{R}$ the sets $X_1 = x_1$ and $X_2 = x_2$ are independent

- Let $X_1, X_2, \ldots, X_n$ independent Bernoulli trials with parameter p
- We construct a random variable as follows:

$$Y = \sum_{i=1}^{n} X_i = X_1 + \cdots + X_n$$

We then say that Y follows a binomial distribution with parameters n and p and write

$$Y \sim binomial(n, p)$$

Exercise 4

- You invite 50 friends to your birthday party. We assume that
  - Friends don't influence each other
  - Each friend has the same chance to attend
- There is a 0.8 chance that any given friend attends the party.

Build our variables:

$X_1, X_2, \dots, X_{50}$ all follow $X_i \sim \text{Bernoulli}(0.8)$

$i = 1, \dots, 50$

$$Y = \sum_{i=1}^{50} X_i \to Y \sim \text{Binomial}(50, 0.8)$$

Chance that exactly 30 friends attend?

$$\boxed{f(30) = P(Y=30) = \binom{50}{30} 0.8^{30} (1-0.8)^{50-30}}$$

PMF of a Binomial distribution

If $Y \sim \textit{binomial}(n, p)$, the PMF of Y is given by

$$\boxed{f(x) = P(Y = x) = \binom{n}{x} p^x (1-p)^{n-x}}$$

for $x \in Y(S)$ and $f(x) = 0$ otherwise.

Expected value

Greek letter "mu"

- Let X be a discrete random variable. The expected value of X is given by

$$\underset{\mu}{\underline{EX}} = \sum_{\underline{x \in X(S)}} x \cdot \overset{P(X=x)}{f}(x)$$

- The expected value can be understood as the long-run average of values that X assigns when the experiment is performed many times

Variance

Greek Letter "sigma"

- Let X be a discrete random variable. The variance of X is given by

$$\overbrace{VarX}^{\sigma^2} = E\left((X-\mu)^2\right) = \sum_{x \in X(S)} (x-\mu)^2 \cdot f(x)$$

- The standard deviation of X is the square root of the variance

$$\sigma = \sqrt{\sigma^2}$$

Properties of the expected value and variance

Let X, Y be random variables and a, b scalars (constant values).

- $E(a + X) = a + EX$
- $E(bX) = bEX$
- $E(X + Y) = EX + EY$

If in addition, X and Y are independent:

- $Var(a + X) = VarX$
- $Var(bX) = b^2 \cdot VarX$
- $Var(X + Y) = VarX + VarY$

Exercise 5

Let X, Y be independent random variables with $EX = 1$, $EY = 2$, $VarX = 4$, $VarY = 9$. Find the expected value and variance of $2X + 1$ and $X - Y$

.