Lecture 6

Discrete Random Variables and Probability Distributions

(DRV)

Recall that a discrete random variable, X, is a function that assigns real values to possible outcomes of a sample space S, and such that the range of X is finite or countably infinite.

A probability distribution for a random variable X is a description of the probabilities associated with the possible values of X.

For a DRV, this is often just a list of the possible outcomes together with probabilities for each. (more complicated situations: a formula).

Ex: Consider the experiment where we toss three fair coins. Let H be the random variable "# of heads".

The range of H is range $\{0, 1, 2, 3\}$, so H is a DVR.

The probability distribution is given by:

$$P(H=0) = P(t,t,t) = 1/8$$
$$P(H=1) = P((t t h), (t h h), (h t t)) = 3/8$$
$$P(H=2) = P((h h t), (h t h), (t h h)) = 3/8$$
$$P(H=3) = P((h h h)) = 1/8.$$

When the math concept is taught in its most basic form in class and the homework requires extra steps

Sometimes, a probability distribution can be conveniently described by a function, called the probability mass function

Definition: Let X be a discrete random variable with range $\{x_1, x_2 \dots\}$. Then a probability mass function is a function f such that:

1) $f(x_i) \geq 0 \quad \forall i$
2) $\sum_i f(x_i) = 1$.
3) $f(x_i) = P(X = x_i)$

Ex: In the coin tossing experiment we can set $f(x) = \frac{\binom{3}{x}}{8}$.

Note: $1/8 + 3/8 + 3/8 + 1/8 = 1$.

Cumulative Distribution functions.

Another way to describe a DVR is via a cumulative distribution function:

Defn: Given a DVR X and probability mass function $f(x)$, the cumulative Distribution Function of X is defined to be

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i).$$

Note that Given a cumulative distribution function $F(x)$ for a DVR, we can recover the probability mass function as

$$f(x_n) = F(x_n) - F(x_{n-1})$$
$$= \sum_{x_i \leq x_n} f(x_i) - \sum_{x_i \leq x_{n-1}} f(x_i)$$

Cummulative distribution functions satisfy the following (for a DVR):

1) $F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$

2) $0 \leq F(x) \leq 1$.

3) $x \leq y \Rightarrow F(x) \leq F(y)$.

<u>Ex:</u> we can have a DVR X with Range $\{1, 2, 3, \ldots\}$ and $P(X = n) = 1/2^n$.

Then $F(n) = \sum_{i=1}^{n} \frac{1}{2^n} = \frac{1}{2}\left(\frac{1 - 1/2^n}{1 - 1/2}\right)$

In particular, $\lim_{n \to \infty} F(n) = 1$.

<u>Mean and Variance of a RRV</u>

- Two numbers used to summarize a probability distribution; do not determine it uniquely.

- Let X be a DVR with prob. mass function $f(x)$.

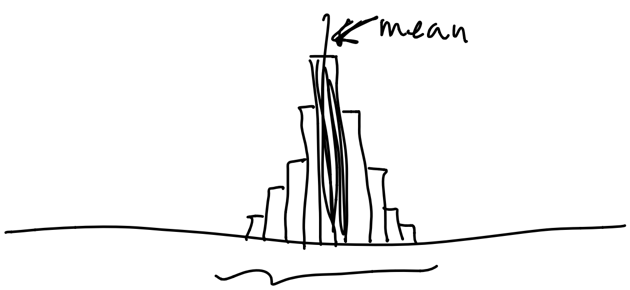
<u>Mean (expected Value)</u> $\ast$ $\mu := E(X) = \sum_x x f(x)$ $\ast$

- is the average of the outcomes, weighted by the probability.

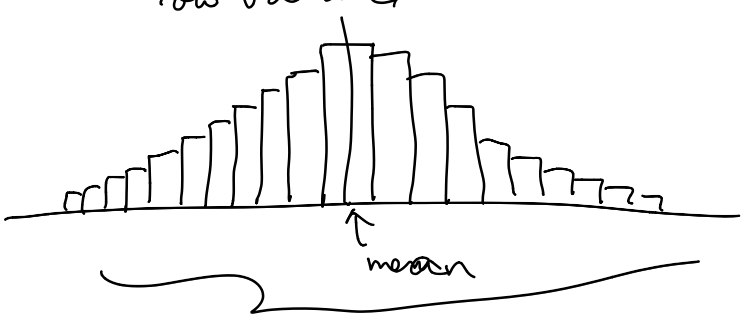
Variance

$$\sigma^2 = V(X) = E(X-\mu)^2 = \sum_x (x-\mu)^2 f(x)$$
$$= \sum_x x^2 f(x) - \mu^2$$

- measures the "spread" of the possible outcomes / how much they vary away from the mean.



low variance



high variance

Ex. Consider the following game:

- a fair die is rolled
- if the roll n is $2, 3, 5$, you win $2 \cdot n$ dollars
- if the roll $n = 1, 4, 6$, you get nothing.

How much would you pay to play one round of this game?

Answer: We consider the expected value of the game! Let W be the discrete random variable "winnings", so $\text{range}(W) = \{0, 2, 4, 6, 10\}$.

$$P(W=0) = 3/6 \quad (\text{for a roll } 1, 4, 6)$$
$$P(W=4) = 1/6 \quad (\text{roll } 2)$$
$$P(W=6) = 1/6 \quad (\text{roll } 3)$$
$$P(W=10) = 1/6 \quad (\text{roll } 5).$$

$$\text{So } \mu = E(W) = \sum f(x)x = \frac{3}{6}0 + 1/6 \cdot 4 + 1/6 \cdot 6 + 1/6 \cdot 10$$
$$= \frac{4 + 6 + 10}{6} = \frac{20}{6}$$
$$= \frac{10}{3} \approx \$3.33.$$

So on average, you should expect to win

\$3.33, so if you pay more than that per round, you should expect to lose money.

(A game where the cost equals the expected value is called a fair game.)

We can also compute the variance of the game:

$$V(X) = \sum_x (x-\mu)^2 f(x)$$

$$= \left(0-\tfrac{10}{3}\right)^2 \cdot \tfrac{3}{6} + \left(4-\tfrac{10}{3}\right)^2 \tfrac{1}{6} + \left(6-\tfrac{10}{3}\right)^2 \tfrac{1}{6} + \left(10-\tfrac{10}{3}\right)^2 \tfrac{1}{6}$$

$$= 14.22$$