4 Discrete Random Variables

Demonstration: Surprize Quiz

correct answers	frequency	relative frequency
0		
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		

Relative frequency histogram:

Definition: A discrete random variable is a function that assigns a number to each outcome of an experiment with a finite sample space.
Notation: A random variable is usually denoted X, and a generic number assigned by X is denoted x.

In the *Surprize Quiz* demonstration, we were interested in the number of correct answers on each quiz. If C stands for correct and I stands for incorrect, an example of an outcome in this experiment is an ordered list like ICIICIIIIC. The random variable should assign to each outcome of the experiment the number that is of interest, namely the number of correct answers. That is, we let $X =$ the number of correct answers. Then $X = 3$ for the outcome ICIICIIIIC.

4.1 Probability Distributions

Discrete random variables are best represented using a table.

Definition: A probability distribution of a discrete random variable X is a table listing the probability of each possible value of X:

x	$P(x)$

Notation: $P(x) = P(X = x)$ is the probability that a randomly chosen outcome will have the value x.

Note that for experiments with equally likely outcomes, $P(x)$ is simply the relative frequency of x.

All discrete random variables have the property

$$\sum P(x) = 1.$$

Example 4.1. Find the probability distribution of $X =$ the number of correct answers in the *Surprize Quiz* demonstration.

Example 4.2. Find the probability distribution of X = the number of heads in 3 coin tosses.

4.2 Expected Value and Standard Deviation

In Section 2.1 we looked at the mean of a data set given as a relative frequency table. The mean of a discrete random variable is defined in the same manner.

Definition: For a discrete random variable, X,

- the mean or expected value of X is

$$\mu = E(X) = \sum xP(x)$$

- the variance of X is

$$\sigma^2 = E(X^2) - \mu^2, \text{ where } E(X^2) = \sum x^2 P(x)$$

- the standard deviation (SD) of X is

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

The SHARP EL-531X calculator can calculate the mean and SD for a probability distribution. Instructions can be found in the Appendix.

Example 4.3. Find the mean and SD of X = the number of correct answers in the *Surprize Quiz* demonstration. Use the probability distribution from Example 4.1.

Since each outcome of an experiment is assigned exactly one number, $P(a \leq X \leq b)$ is simply the sum of all the probabilities for x-values that fall within $[a, b]$.

Example 4.4. Given the following probability distribution

x	$P(x)$
-5	0.15
-2	0.2
1	0.4
6	0.25

find:

(a) $P(-2.5 \leq X \leq 2.5)$.

(b) the mean of X.

(c) the variance of X.

(d) the standard deviation of X.

(e) the probability that an x-value lies within one standard deviation of the mean; i.e. $P(\mu - \sigma \leq X \leq \mu + \sigma)$.

Example 4.5. Project 1 has a 35% chance of earning \$0, a 50% chance of earning \$300,000, and a 15% chance of earning \$800,000.

Project 2 has a 60% chance of earning \$0 and a 40% chance of earning \$1,000,000.

(a) Find the probability distributions of the earnings for each project.

(b) Find the expected earnings for each project.

(c) Find the standard deviation of earnings for each project.

(d) Which project has higher expected earnings?

(e) In terms of earnings, which project is riskier?

Example 4.6. Suppose you want to insure a \$2,000 tablet against theft for one year by paying a premium m, and that the probability of theft is 4.7%.

(a) Find the probability distribution of the insurance company's gain.

(b) Find the premium (i.e. the value of m) if the insurance company expects to gain \$40.

Additional Notes

Additional Notes