Random variables

random variable is a mapping from the sample space to real numbers. notation: X, Y, Z,...
Example: Ask a student whether she/he works part time or not.
$\mathcal{S}= \{Yes, No\}$.
X=1 if Yes, X=0 if No.
Y = number of car accidents in NYC in a week.
Flip a coin three times. Let Z=number of heads in three flips.
$\{TTT\} \rightarrow Z = 0$
$\{HTT, THT, TTH\} \rightarrow Z = 1$
$\{HHT, HTH, THH\} \rightarrow Z = 2$
$\{HHH\} \rightarrow Z = 3$.
W = the weight of a randomly selected athlete.

Discrete and continuous random variables

A **discrete random variable** takes finite or countably infinite values.
A **continuous random variable** takes all values in an interval or union of intervals.

The Binomial distribution

Binomial experiment:

1. The experiment consists of n trials.
2. Each trial has two possible outcomes, success or failure.
3. The trials are independent.
4. The probability of success, p, is constant from trial to trial.

Binomial random variable X is defined to be the number of successes out of n trials. Note the possible vales X can take are 0, 1, 2, ...n.

$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, k = 0, 1, \cdots, n.$

Computing binomial probabilities

An IT center uses 9 aging disk drives for storage. The probability that any one of them is out of service is 0.06. For the center to function properly, at least 7 of the drives must be available. What is the probability that the center can get its work done?
Let X = number of drives available, then
$P(X \geq 7) = P(X = 7) + P(X = 8) + P(X = 9) =$
$\binom{9}{7}0.94^7 0.06^2 + \binom{9}{8}0.94^8 0.06^1 + \binom{9}{9}0.94^9 .06^0 = 0.986.$

Exercise

An analyst has tracked a stock for the past six months and found that on any given day, it either goes up a point or goes down a point. Furthermore, it went up on 25 percent of the days and down on 75 percent. What the probability that at the close of trading 4 days from now, the price of the stock will be the same as it is today? Assume the daily movements are independent.

solutions

The number of days for which the stock rises is binomial with
$n = 4, p = 0.25$.
For $k = 2$, we have $\binom{4}{2} 0.25^2 0.75^2 = 0.211$.

The hypergeometric distribution

Sampling without replacement:
Suppose an urn contains r red chips and N-r white chips. If n chips are drawn at random, without replacement and let $X =$ the number of red chips selected. Then
$P(X = k) = \binom{r}{k}\binom{N-r}{n-k}/\binom{N}{n}$ for $k = 0, 1, 2, \cdots, n, k \leq r$, and $n - k \leq N - r$.
example: An inspector decides to examine the exhaust of six of a company's 24 trucks. If four of the company's trucks emit excessive amounts of pollutants, what the probability that none of them will be included in the inspector's sample?
$P(X = 0) = \binom{4}{0}\binom{20}{6}/\binom{24}{6} = 0.288.$

exercise

Among the 25 applicants for a job, ten have college degrees. If 4 of the applicants are randomly chosen for an interview, what is the probability that two has a college degree?

$$P(X = 2) = \frac{\binom{10}{2}\binom{15}{2}}{\binom{25}{4}} = 0.3735.$$

Discrete random variables

If X is a discrete random variable, the function given by $f(x) = P(X = x)$ for each x within the range of X is called the probability distribution or probability mass function (pmf) of X. A pmf satisfies

1. $f(x) \geq 0$ and $\sum_x f(x) = 1$.

example: check the function given by
$f(x) = \frac{x+2}{25}$ for $x = 1, 2, 3, 4, 5$ can serve as a pmf.
Solution: each $f(x) \geq 0$, and $\sum f(x) = 1$.

cumulative distribution

If X is a discrete random variable, the function given by $F(x) = P(X \leq x) = \sum_{t \leq x} f(t), -\infty \leq x\infty$ is called the cumulative distribution function (cdf) of X.

Example : The pmf of a rv X is given below:

x	1	2	3	4
p(x)	0.4	0.3	0.2	0.1

Note that

$F(1) = P(X \leq 1) = P(X = 1) = 0.4$

$F(2) = P(X \leq 2) = P(X = 1or2) = 0.4 + 0.3 = 0.7$

$F(3) = P(X \leq 3) = P(X = 1or2or3) = 0.4 + 0.3 + 0.2 = 0.9$

$F(4) = P(X \leq 4) = P(X = 1or2or3or4) = 0.4 + 0.3 + 0.2 + 0.1 = 1.0$

The cdf of X is
$F(x) = 0$ if $x < 1$
$=0.4$ if $1 \leq x < 2$
$=0.7$ if $2 \leq x < 3$
$=0.9$ if $3 \leq x < 4$
$=1.0$ if $x \geq 4$.

Exercise

3.3.14: At the points $x = 0, 1, 2, \cdots, 6$, the cdf for the discrete random variable X has the value $F_X(x) = x(x+1)/42$. Find the pmf of X.