Business Analytics

Discrete Random Variables

Faton Berisha

Chapter 4

Discrete Random Variables

Discrete Random Variables

Random Variables

A *random variable* is a variable that assumes numerical values that are determined by the outcome of an experiment

Discrete random variable: Possible values can be counted or listed

- E.g., the number, *x*, of defective units in a batch of 20; a listener rating (1 to 5) in an AccuRating music survey; number of major fires in Tetovo during the last year.
- Example 4.2. The number of correct answers to three pop quiz questions

Random Variables (Cont.)

Continuous random variable: May assume any numerical value in one or more intervals

- E.g., the waiting time, *x*, for a credit card authorization; the interest rate charged on a business loan
- Example 4.2. The car mileage case

Discrete Probability Distributions

The *probability distribution* of a discrete random variable is a table, graph or formula that gives the probability associated with each possible value that the variable can assume

Notation: Denote the values of the random variable by x and the value's associated probability by $p(x)$

Properties

1. For any value x of the random variable, $p(x) \geq 0$
2. The probabilities of all the events in the sample space must sum to 1, that is,

$$\sum_{\text{all } x} p(x) = 1$$

Example: Number of Correct Answers

Example 4.2. The number of correct answers of a student to three questions in a pop quiz

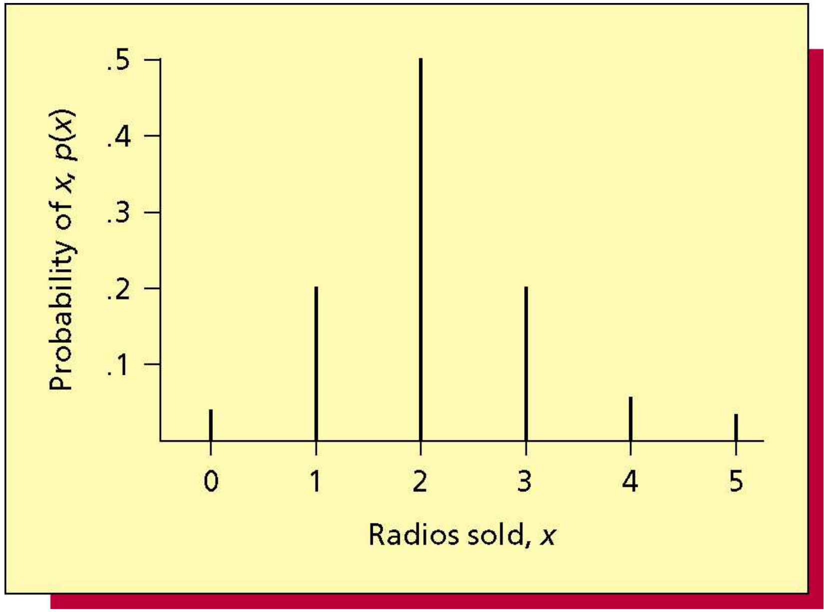
Assume that the student studiesand has a 0.9 probability of answering each question correctly.

x, Correct answers	p(x), Probability
0	p(0) = 0.001
1	p(1) = 0.027
2	p(2) = 0.243
3	p(3) = 0.729

Example: Number of Radios Sold

Example 4.3. Number of Radios Sold at Sound City in a Week

x, Radios	p(x), Probability
0	p(0) = 0.03
1	p(1) = 0.20
2	p(2) = 0.50
3	p(3) = 0.20
4	p(4) = 0.05
5	p(5) = 0.02



Expected Value of a Discrete Random Variable

The *mean* or *expected value* of a discrete random variable X is:

$$\mu_x = \sum_{\text{All } x} x p(x)$$

μ_x is the value expected to occur in the long run and on average.

Example 4.4: Expected Number of Radios Sold in a Week

x, Radios	$p(x)$, Probability	$x\,p(x)$
0	p(0) = 0.03	0(0.03) = 0.00
1	p(1) = 0.20	1(0.20) = 0.20
2	p(2) = 0.50	2(0.50) = 1.00
3	p(3) = 0.20	3(0.20) = 0.60
4	p(4) = 0.05	4(0.05) = 0.20
5	p(5) = 0.02	5(0.02) = 0.10
	1.00	2.10

Variance and Standard Deviation

The *variance* of a discrete random variable is:

$$\sigma_X^2 = \sum_{All\ x} (x - \mu_X)^2 p(x)$$

❖The variance is the average of the squared deviations of the different values of the random variable from the expected value

The *standard deviation* is the square root of the variance

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

❖The variance and standard deviation measure the spread of the values of the random variable from their expected value

Example: Variance and Standard Deviation

Example 4.7: Variance and Standard Deviation of the Number of Radios Sold in a Week

x, Radios	$p(x)$, Probability	$(x - \mu_x)^2 \; p(x)$
0	$p(0) = 0.03$	$(0 - 2.1)^2 (0.03) = 0.1323$
1	$p(1) = 0.20$	$(1 - 2.1)^2 (0.20) = 0.2420$
2	$p(2) = 0.50$	$(2 - 2.1)^2 (0.50) = 0.0050$
3	$p(3) = 0.20$	$(3 - 2.1)^2 (0.20) = 0.1620$
4	$p(4) = 0.05$	$(4 - 2.1)^2 (0.05) = 0.1805$
5	$p(5) = 0.02$	$(5 - 2.1)^2 (0.02) = 0.1682$
	1.00	0.8900

Variance

$$\sigma_X^2 = 0.89$$

Standard deviation

$$\sigma_X = \sqrt{0.89} = 0.9434$$

The Binomial Distribution

The Binomial Experiment (identical trials):

1. Experiment consists of *n* identical trials
2. Each trial results in either "success" or "failure"
3. Probability of success, *p*, is constant from trial to trial
4. Trials are independent

Note: The probability of failure, *q*, is 1 – *p* and is constant from trial to trial

If *x* is the total number of successes in *n* trials of a binomial experiment, then *x* is ***a binomial random variable***

The Binomial Distribution (Cont.)

For a binomial random variable *x*, the probability of *x* successes in *n* trials is given by the *binomial distribution*:

$$p(x) = \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

❖Note: *n*! is read as “*n* factorial” and *n*! = *n* (*n*-1) (*n*-2) ... 1

❖ E.g., 5! = 5・4・3・2・1 = 120

❖Also, 0! =1

❖Factorials are not defined for negative numbers or fractions

Example: Side effect of a drug

Shembulli 4.10. Nausea as a side effect experienced by patients treated with Phe-Mycin

- x = number of patients who will experience nausea following treatment with Phe-Mycin.
- Probability that 2 of the 4 patients treated will experience nausea

$$n = 4,\ p = 0.1,\ q = 1 - p = 1 - 0.1 = 0.9$$

$$p(2) = \mathrm{P}(x = 2) = \frac{4!}{2!(4-2)!}(0.1)^2(0.9)^{4-2} = 6(0.1)^2(0.9)^{4-2} = 0.0486$$

Example: Binomial Distribution, n = 4, p = 0.1

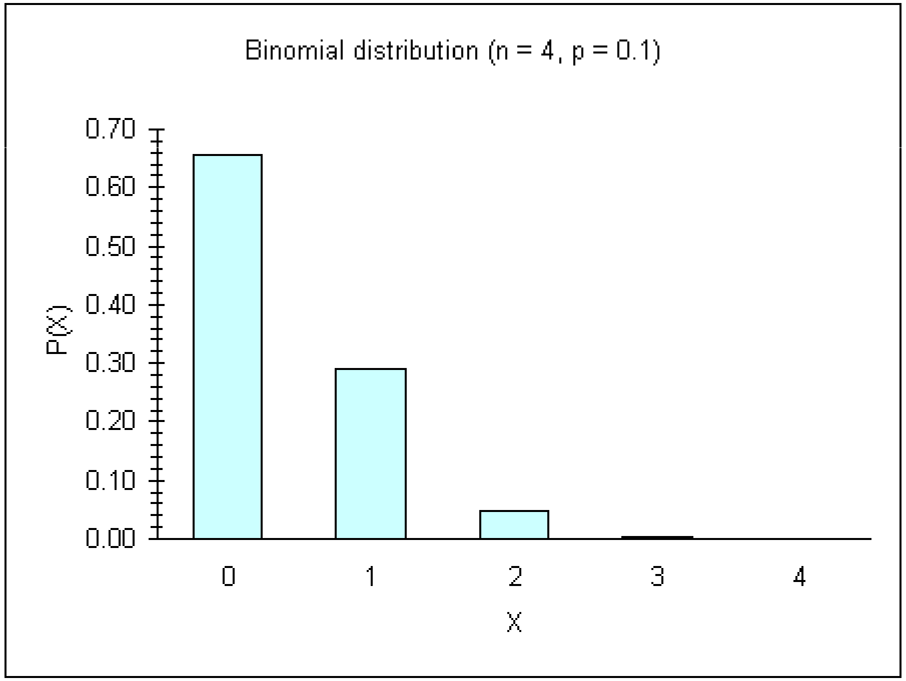
(a) MegaStat output of the binomial distribution

Binomial distribution

4 n
0.1 p

X	*P(X)*	*cumulative probability*
0	0.65610	0.65610
1	0.29160	0.94770
2	0.04860	0.99630
3	0.00360	0.99990
4	0.00010	1.00000
	1.00000	

0.400 expected value
0.360 variance
0.600 standard deviation



Example: Side effect of a drug (Cont.)

- Example 4.11. Investigating the claim that $p = 0.10$.
  - Suppose that in the sample of $n = 4$, 3 patients experienced nausea.
  - $P(x \geq 3) = P(x = 3) + P(x = 4)$
    $= 0.0036 + 0.0001 = 0.0037$
  - If $p = 0.1$, then in 0.37% of all possible samples of 4 randomly selected patients, at least 3 of the patients experience nausea.

Example: Quality Control

Example 4.12. investigating the claim that $p = 0.95$

- A TV set producer claims that $p = 0.95$ is the probability that its sets last at least 5 years without requiring a single repair.
- Suppose that out of a random sample of size n = 8, 5 sets have lasted at least 5 years without a single repair.
- $P(x \leq 5) = . . . = 0.0058$

Binomial Probability Table

Table 4.7(a) for $n = 4$, with $x = 2$ and $p = 0.1$

$p = 0.1$

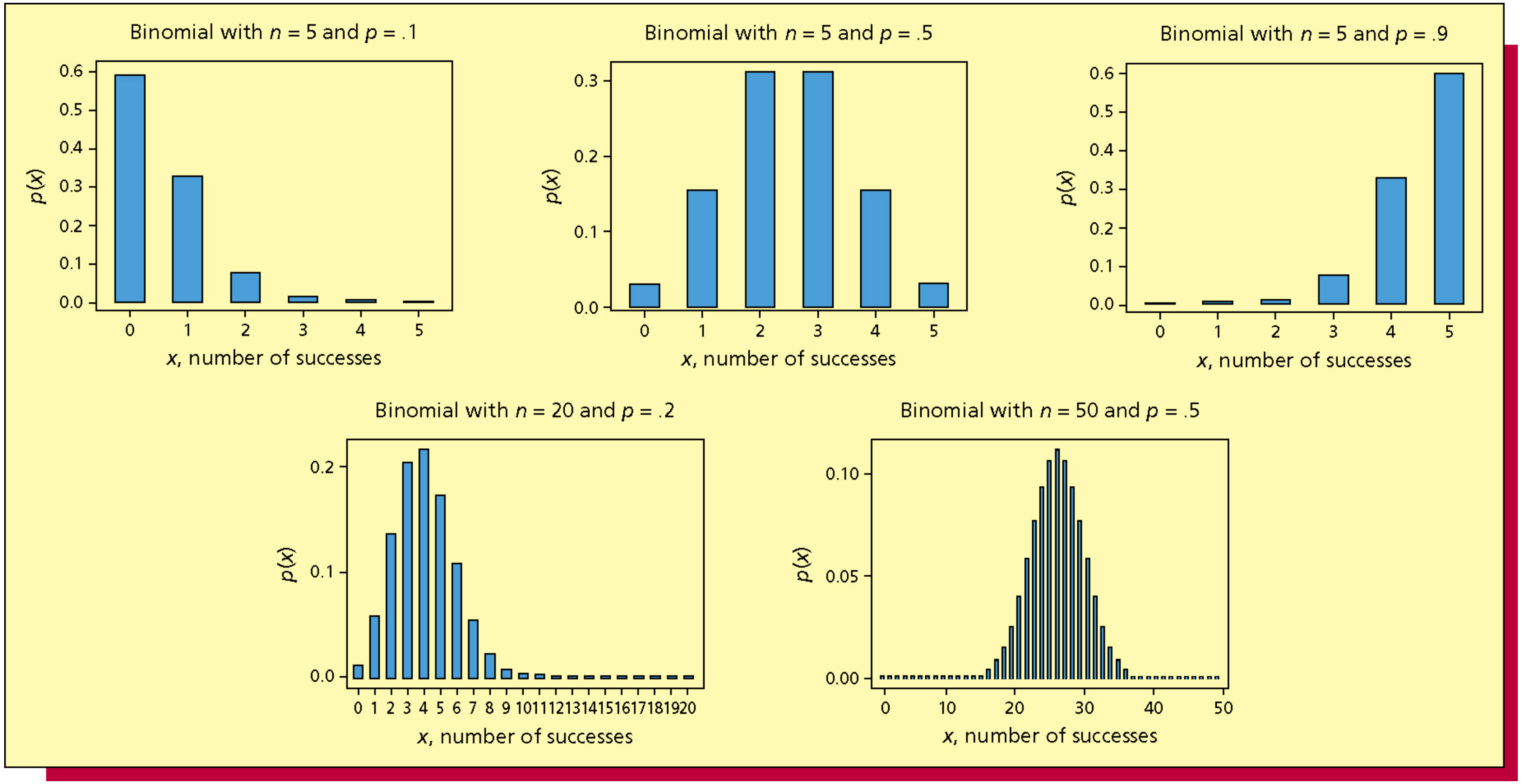
Values of p (.05 to .50)

x	0.05	0.1	0.15	…	0.50	
0	0.8145	0.6561	0.5220	…	0.0625	4
1	0.1715	0.2916	0.3685	…	0.2500	3
2	0.0135	0.0486	0.0975	…	0.3750	2
3	0.0005	0.0036	0.0115	…	0.2500	1
4	0.0000	0.0001	0.0005	…	0.0625	0
	0.95	0.9	0.85	…	0.50	x

Values of p (.05 to .50)

$P(x = 2) = 0.0486$

Several Binomial Distributions



Mean and Variance of a Binomial Random Variable

If x is a binomial random variable with **parameters** n and p (so $q = 1 – p$), then

Mean $\mu_X = np$

Variance $\sigma_x^2 = npq \,, \;\; q = 1 - p$

Standard deviation $\sigma_{\mathrm{x}} = \sqrt{npq}$

The Poisson Distribution

Consider the number of times an event occurs over an interval of time or space, and assume that

1. The probability of occurrence is the same for any intervals of equal length
2. The occurrence in any interval is independent of an occurrence in any non-overlapping interval

If x = the number of occurrences in a specified interval, then x is a *Poisson random variable*

The Poisson Distribution (Cont.)

Suppose μ is the mean or expected number of occurrences during a specified interval

The probability of *x* occurrences in the interval when μ are expected is described by the *Poisson distribution*:

$$p(x) = \frac{e^{-\mu}\mu^{x}}{x!}$$

where *x* can take any of the values *x* = 0, 1, 2, 3, …

and *e* = 2.71828… (*e* is the base of the natural logs)

Example: Air Traffic Control

Example 4.13.

x = the number of air traffic control errors in a week in an airport

$\mu = 0.4$ (the expected number of errors in a week)

Find the probability of 3 errors occurring in a week.

$$p(3) = P(x = 3) = \frac{e^{-0.4}(0.4)^3}{3!} = 0.0072$$

Poisson Probability Table

Table 4.9

μ=0.
4

μ, mean number of
Occurrences

x	0.1	0.2	...	0.4	...	1.00
0	0.9048	0.8187	...	0.6703	...	0.3679
1	0.0905	0.1637	...	0.2681	...	0.3679
2	0.0045	0.0164	...	0.0536	...	0.1839
3	0.0002	0.0011	...	0.0072	...	0.0613
4	0.0000	0.0001	...	0.0007	...	0.0153
5	0.0000	0.0000	...	0.0001	...	0.0031

$$p(x=3)=\frac{e^{-0.4}(0.4)^3}{3!}=0.0072$$

Poisson Probability Calculations

x, the Number of Errors in a Week	$p(x) = \dfrac{e^{-\mu}\mu^x}{x!}$
0	$p(0) = \dfrac{e^{-.4}(.4)^0}{0!} = .6703$
1	$p(1) = \dfrac{e^{-.4}(.4)^1}{1!} = .2681$
2	$p(2) = \dfrac{e^{-.4}(.4)^2}{2!} = .0536$
3	$p(3) = \dfrac{e^{-.4}(.4)^3}{3!} = .0072$
4	$p(4) = \dfrac{e^{-.4}(.4)^4}{4!} = .0007$
5	$p(5) = \dfrac{e^{-.4}(.4)^5}{5!} = .0001$
6	$p(6) = \dfrac{e^{-.4}(.4)^6}{6!} = .0000$

Mean and Variance of a Poisson Random Variable

If x is a Poisson random variable with parameter μ, then

Mean $\mu_X = \mu$

Variance $\sigma_x^2 = \mu$

Standard deviation $\sigma_x = \sqrt{\sigma_x^2} = \sqrt{\mu}$

Example: Poisson Distribution, $\mu = 0.4$

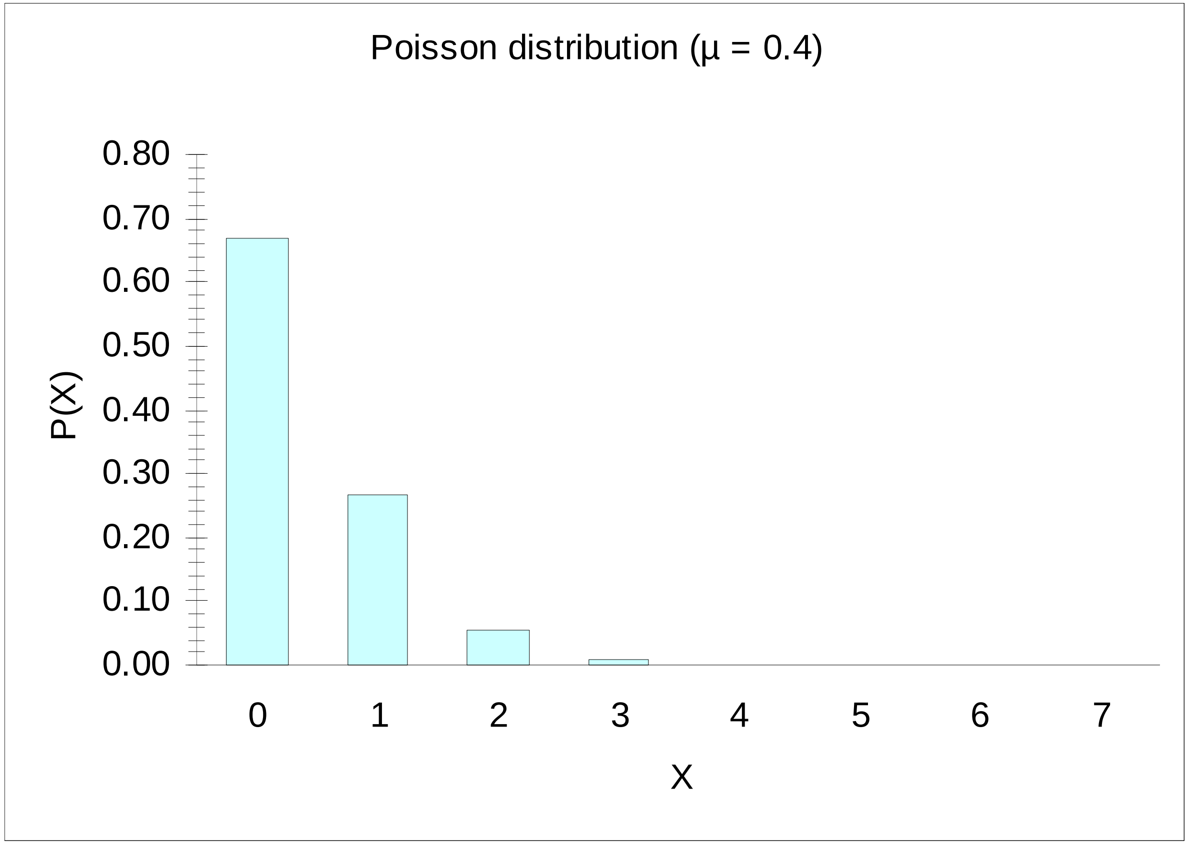
(a) MegaStat output on Poisson distribution

Poisson distribution

0.4 mean rate o

X	P(X)	cumulative probability
0	0.67032	0.67032
1	0.26813	0.93845
2	0.05363	0.99207
3	0.00715	0.99922
4	0.00072	0.99994
5	0.00006	1.00000
6	0.00000	1.00000
7	0.00000	1.00000
8	0.00000	1.00000
	1.00000	

0.400 expected va
0.400 variance
0.632 standard de



Several Poisson Distributions