Chapter 5: Discrete Probability Distributions

5.1 Random Variable

Random Variable: numerical description of the outcome of an experiment.

Discrete Random Variable: may assume either a finite number of values or an infinite sequence of values such as 0, 1, 2, …

e.g.: x = number of parts of the CPA exam passed.
e.g.: x = number of cars arriving at a tollbooth during a one-day period.
e.g.: x = 0 if the individual can recall the message in a television commercial and x = 1 if s/he cannot.

Continuous Random Variable: may assume any numerical value in an interval or collection of intervals.

e.g.: x = time between consecutive incoming calls to an insurance company claims office, in minutes.
e.g.: x = number of miles to the location of the next traffic accident along a 90-mile section of I-75.

5.2 Developing Discrete Probability Distributions

Probability distribution: describes how probabilities are distributed over the values of the random variable.

Discrete Probability function, f(x): provides the probability for each value of the random variable
$f(x) \geq 0$ and $\Sigma f(x) = 1$.

e.g.: x = number obtained on one roll of a die; f(x) = the probability of x
Table 5.3 Probability distribution for number obtained on one roll of a die

x	1	2	3	4	5	6
f(x)	1/6	1/6	1/6	1/6	1/6	1/6

Empirical discrete distribution: obtained when the relative frequency method (see Ch.4.1) is used to develop the discrete probability distribution.

Discrete uniform probability function: f(x) = 1/n, where n = # of values the random variable may assume

5.3 Expected Value and Variance

Expected Value of a discrete random variable: $E(x) = \mu = \Sigma x f(x)$
It is a weighted average of the values of the random variable, where the weights are the probabilities. The expected value does not have to be a value the random variable can assume.

Variance of a discrete random variable: $Var(x) = \sigma^2 = \Sigma(x - \mu)^2 f(x)$
It is a weighted average of the squared deviations of a random variable from its mean, where the weights are the probabilities.

Standard Deviation of a discrete random variable: $\sigma = \sqrt{\sigma^2}$
It is measured in the same units as the random variable, while σ^2 is measured in squared units.

5.4 Bivariate Distributions, Covariance, and Financial Portfolios

Bivariate probability distributions: probability distribution involving two random variables.

Covariance of random variables X and Y:
$Cov(x,y) = \sigma_{xy} = \Sigma(x_i - \mu_x)(y_j - \mu_y)f(x_i, y_j)$
$Cov(x,y) = \sigma_{xy} = E[(x_i - \mu_x)(y_j - \mu_y)]$
$Cov(x,y) = \sigma_{xy} = [Var(x + y) - Var(x) - Var(y)]/2$
$Cov(x,y) = \sigma_{xy} = E(xy) - \mu_x\mu_y$

Recall:
(Ch. 3.1) Mean:
$\mu = \Sigma x_i/N$, for a population
NOTE: The mean is computed by giving each observation a weight of 1/N.
(Ch. 3.2) Variance:
$\sigma^2 = (\Sigma(x_i - \mu)^2)/N$, for a population
NOTE: The variance is the average of the squared deviations from the mean, or the weighted average, where 1/N is the weight for each squared deviation.
(Ch. 3.5) Covariance:
$\sigma_{xy} = (\Sigma(x_i - \mu_x)(y_i - \mu_y))/N$, for a population
NOTE: The covariance is the average of the products of the deviation of x from its mean and the deviation y from its mean, or the weighted average, where 1/N is the weight for each product of the deviation of x from its mean and the deviation of y from its mean.

Now, looking at discrete probability distributions:
(Ch. 5.3) Expected Value of a discrete random variable:
$E(x) = \mu_x = \Sigma xf(x)$
NOTE: The expected value of a discrete random variable x is a weighted average of the values of the random variable, where the weights are the probabilities of occurrence of each value x_i.
(Ch. 5.3) Variance of a discrete random variable:
$Var(x) = \sigma^2 = \Sigma(x_i - \mu_x)^2f(x_i)$
NOTE: The variance of a discrete random variable x is a weighted average of the squared deviations of the random variable from its mean, where the weights are the probabilities of occurrence of each value x_i. IN OTHER WORDS, the variance of a discrete random variable x is the expected value of the squared deviations of that random variable from its mean:
$Var(x) = E(x - \mu_x)^2$
(Ch. 5.4) Covariance of discrete random variables x and y:
$Cov(x,y) = \sigma_{xy} = \Sigma_{ij}(x_i - \mu_x)(y_j - \mu_y)f(x_i, y_j)$
NOTE: The covariance of discrete random variables x and y is a weighted average of the products of the deviation of random variable x from its mean and the deviation of random variable y from its mean, where the weights are the probabilities of joint occurrence of values (x_i, y_j), or joint probabilities. IN OTHER WORDS, the covariance of discrete random variables x and y is the expected value of the products of the deviation of random variable x from its mean and deviation of random variable y from its mean:
$Cov(x,y) = E[(x_i - \mu_x)(y_j - \mu_y)]$

Similarly, we define Var(x+y) as a expected value of the squared deviations of (x+y) from its mean:
$Var(x+y) = E[(x+y) - (\mu_x+\mu_y)]^2$
This leads to :
$Var(x+y) = E[(x+y)^2 + (\mu_x+\mu_y)^2 - 2(x+y)(\mu_x+\mu_y)]$
$Var(x+y) = E[x^2 + y^2 + 2xy + \mu_x^2 + \mu_y^2 + 2\mu_x\mu_y - 2x\mu_x - 2x\mu_y - 2y\mu_x - 2y\mu_y]$
$Var(x+y) = E[(x^2 + \mu_x^2 - 2x\mu_x) + (y^2 + \mu_y^2 - 2y\mu_y) + 2(xy + \mu_x\mu_y - x\mu_y - y\mu_x)]$
$Var(x+y) = E[(x - \mu_x)^2 + (y - \mu_y)^2 + 2(x - \mu_x)(y - \mu_y)]$
$Var(x+y) = E(x - \mu_x)^2 + E(y - \mu_y)^2 + 2E[(x - \mu_x)(y - \mu_y)]$
$Var(x+y) = Var(x) + Var(y) + 2Cov(x,y)$
Cov(x,y) is thus equal to:
$Cov(x,y) = [Var(x+y) - Var(x) - Var(y)]/2$

Alternative formulas for Var and Cov:
$Var(x) = E(x - \mu_x)^2 = Var(x) = E(x^2 + \mu_x^2 - 2x\mu_x) = E(x^2) + \mu_x^2 - 2\mu_xE(x) = E(x^2) + \mu_x^2 - 2\ \mu_x\mu_x = E(x^2) + \mu_x^2 - 2\mu_x^2$
$Var(x) = E(x^2) - \mu_x^2$

$Cov(x,y) = E[(x_i - \mu_x)(y_j - \mu_y)]$
$Cov(x,y) = E(xy - \mu_xy - x\mu_y + \mu_x\mu_y) = E(xy) - \mu_xE(y) - \mu_yE(x) + \mu_x\mu_y = E(xy) - \mu_x\mu_y - \cancel{\mu_y}\mu_x + \cancel{\mu_x}\mu_y$
$Cov(x,y) = E(xy) - \mu_x\mu_y$

Correlation between random variables X and Y (*see also 3.5*):
$\rho_{xy} = \sigma_{xy}/\sigma_x\sigma_y$

Expected value of a linear combination of random variables X and Y:
$E(ax + by) = aE(x) + bE(y)$

Proof:
$E(ax + by) = \Sigma_{ij}(ax_i + by_j)f(x_i, y_j) = \Sigma_{ij}ax_if(x_i, y_j) + \Sigma_{ij}by_jf(x_i, y_j) = a\Sigma_{ij}x_if(x_i, y_j) + b\Sigma_{ij}y_jf(x_i, y_j)$
$E(ax + by) = a\Sigma_ix_i\Sigma_jf(x_i, y_j) + b\Sigma_jy_j\Sigma_if(x_i, y_j) = a\Sigma_ix_if(x_i) + b\Sigma_jy_jf(y_j) = aE(x) + bE(y)$

Variance of a linear combination of two random variables:
$Var(ax + by) = a^2Var(x) + b^2Var(y) + 2abCov(x,y)$

Proof:
$Var(ax+by) = E[(ax+by) - (a\mu_x+b\mu_y)]^2 = E[(ax+by)^2 + (a\mu_x+b\mu_y)^2 - 2(ax+by)(a\mu_x+b\mu_y)]$
$Var(ax+by) = E[a^2x^2 + b^2y^2 + 2abxy + a^2\mu_x^2 + b^2\mu_y^2 + 2ab\mu_x\mu_y - 2a^2x\mu_x - 2abx\mu_y - 2aby\mu_x - 2b^2y\mu_y]$
$Var(ax+by) = E[(a^2x^2 + a^2\mu_x^2 - 2\ a^2x\mu_x) + (b^2y^2 + b^2\mu_y^2 - 2\ b^2y\mu_y) + 2ab(xy + \mu_x\mu_y - x\mu_y - y\mu_x)]$
$Var(ax+by) = E[a^2(x - \mu_x)^2 + b^2(y - \mu_y)^2 + 2ab(x - \mu_x)(y - \mu_y)]$
$Var(ax+by) = a^2E(x - \mu_x)^2 + b^2E(y - \mu_y)^2 + 2abE[(x - \mu_x)(y - \mu_y)]$
$Var(ax+by) = a^2Var(x) + b^2Var(y) + 2abCov(x,y)$

Table 5.7 Number of automobiles sold at Dicarlo's Saratogo and Geneva dealerships over 300 days, is the same as Table 4.4 or Table 2.10: crosstabulation of supplier & part quality.
Table 5.8 **Bivariate empirical discrete probability distribution** for sales in Saratoga & Geneva, is the same as the **joint probability** Table 4.5 or Table 2.10b: **crosstabulation of relative frequencies** for supplier & part quality (Ch.2.3 Presentation).

Consider daily sales at DiCarlo Motors, which include sales at its Saratoga, New York automobile dealership and sales at its Geneva, New York automobile dealership.

Table 5.7 Crosstabulation (frequencies)/# of cars sold at two dealerships: (x,y) over 300 days.

		y: # cars sold at the Saratoga Dealership						
		0	**1**	**2**	**3**	**4**	**5**	TOTAL: x Frequency Distribution
x: # cars sold at the Geneva Dealership	**0**	21	30	24	9	2	0	86
	1	21	36	33	18	2	1	111
	2	9	42	9	12	3	2	77
	3	3	9	6	3	5	0	26
	TOTAL: y Frequency Distribution	54	117	72	42	12	3	300

We use Excel as we did in Ch.2.3 Presentation to go from Table 5.7 above to Table 5.8 below.

Table 5.8 Crosstabulation (relative frequencies)/Joint or Bivariate probability distribution: f(x,y)

		y: # cars sold at the Saratoga Dealership						
		0	**1**	**2**	**3**	**4**	**5**	TOTAL: f(x)
x: # cars sold at the Geneva Dealership	**0**	0.07	0.10	0.08	0.03	0.01	0.00	0.29
	1	0.07	0.12	0.11	0.06	0.01	0.00	0.37
	2	0.03	0.14	0.03	0.04	0.01	0.01	0.26
	3	0.01	0.03	0.02	0.01	0.02	0.00	0.09
	TOTAL: f(y)	0.18	0.39	0.24	0.14	0.04	0.01	1.00

Body of the table: $f(x_i,y_j)$, joint or bivariate probabilities
Bottom margin: $f(y_j) = \Sigma_i f(x_i,y_j)$, marginal probabilities or y relative frequency distribution
Right margin: $f(x_i) = \Sigma_j f(x_i,y_j)$, marginal probabilities or x relative frequency distribution

Let s be the total daily sales at DiCarlo Motors: $s = x + y$
$E(s) = E(x+y) = E(x) + E(y)$,
where $E(x) = \Sigma xf(x)$ and $E(y) = \Sigma yf(y)$.
$Var(s) = Var(x+y) = Var(x) + Var(y) + 2Cov(x,y)$,
where $Var(x) = \Sigma(x_i - \mu_x)^2 f(x_i)$, $Var(y) = \Sigma(y_j - \mu_y)^2 f(y_j)$ and $Cov(x,y) = \Sigma_{ij}(x_i - \mu_x)(y_j - \mu_y)f(x_i, y_j)$.

$E(s) = 2.6433$
$Var(s) = 2.3895$
See Ch. 5 DiCarlo Motors Excel file.

5.5 Binomial Probability Distribution

Bernoulli process:
Consists of a trial whose outcome can be classified as either a *success* (with probability p) or a *failure* (with probability 1 – *p*).

Binomial Experiment:
Consists of a sequence of *n* identical and independent trials, each with two outcomes: *success* (with probability p) or *failure* (with probability 1 – *p*).

Stationarity assumption:
The property of the binomial experiment, that the probability of a success, *p*, and thus, the probability of a success, 1 −p, do not change from trial to trial, is referred as the stationarity assumption. Note, it is not to be confused with the independence of trials property of the binomial experiment.

Binomial probability function (*Microsoft Excel function*: **binom.dist)**:

$$f(x) = \binom{n}{x} p^x (1-p)^{(n-x)} \qquad (1)$$

where
f(x) = probability of x successes in n trails
p = probability of a success on one trial

$$\binom{n}{x} = \frac{n!}{x!\,(n-x)!}$$

NOTE: The probability function is applicable for values x = 0, 1,2, …, *n*.

The validity of (1) stems from applying the counting rules for multiple-step experiments and combinations (Ch.4.1). By the counting rule for multiple-step experiments, the probability of any particular sequence of n outcomes containing x successes and n – x failure is p on the first step, p on the second step, …, p on the x^{th} step, (1 – p) on the $(x+1)^{th}$ step, (1 – p) on the $(x+2)^{th}$ step, …, (1 – p) on the n^{th} step: $p^x(1-p)^{n-x}$. Equation (1) then follows, since there are $\binom{n}{x}$ different sequences of the n outcomes leading to x successes and n – x failures (see Ch.4.6 in Ross, 9th).

Expected value and variance for the binomial distribution:
E(x) = μ = *np*
Var(x) = σ^2 = *np*(1 − *p*)

Proofs: see pp. 131-2, Ch.4.6.1 in Ross, 9^{th}.

5.6 Poisson Probability Distribution

Poisson Experiment:
Consists of independent occurrences, over intervals of equal length, each with *constant* probability for an interval of given length.

Poisson probability function (*Microsoft Excel function*: poisson.dist):
f(x) = $\frac{\mu^x e^{-\mu}}{x!}$
where
f(x) = probability of x occurrences in an interval
μ = expected value or mean number of occurrences in an interval
e = 2.71828
NOTE: The probability function is applicable for values x = 0, 1,2, … without limit.

Expected value and variance for the Poisson distribution:
μ = var(x)

Proofs: see pp. 131-8, Ch.4.7 in Ross, 9th.

5.7 Hypergeometric Probability Distribution

Hypergeometric Experiment:
Consists of a sequence of *n*, non-independent trials, each with two outcomes: *success* (with probability p) and *failure* (with probability 1 – *p*), where the probability *p* changes from trial to trial.

Hypergeometric probability function (*Microsoft Excel function*: hypgeom.dist):
f(x) = $\frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}$
where
f(x) = probability of x successes in n trails
N = number of elements in the population
r = number of elements in the population labeled success
NOTE: The probability function is applicable for values x = 0, 1,2, …, *n* where x ≤ r and n – x ≤ N – r.

Expected value and variance for the hypergeometric probability distribution:
E(x) = μ = *n*(*r*/*N*)
Var(x) = σ^2 = *n*(*r*/*N*)(1 – (*r*/*N*))((*N* − *n*)/(*N* − 1))
where
p = (*r*/*N*) is the probability of a success on the first trial
NOTE: If the population size is large, the term (*N* − *n*)/(*N* − 1) approaches 1; E(x) and Var(x) then become: E(x) = μ = *np* & Var(x) = σ^2 = *np*(1 – *p*), as the binomial distribution's.