# Lecture 3

# 3.1. Discrete Probability Distributions

## *Summary*

A random variable provides a numerical description of the outcome of an experiment. The probability distribution for a random variable describes how the probabilities are distributed over the values the random variable can assume. For any discrete random variable *x*, the probability distribution is defined by a probability function, denoted by $f(x)$, which provides the probability associated with each value of the random variable. Once the probability function is defined, we can compute the expected value, variance, and standard deviation for the random variable.

The **binomial distribution** can be used to determine the probability of x successes in n trials whenever the experiment has the following properties:

1. The experiment consists of a sequence of *n* identical trials.

2. Two outcomes are possible on each trial, one called success and the other failure.

3. The probability of a success *p* does not change from trial to trial. Consequently, the probability of failure, $1–p$, does not change from trial to trial.

4. The trials are independent.

When the four properties hold, the binomial probability function can be used to determine the probability of obtaining *x* successes in *n* trials. Formulas were also presented for the mean and variance of the binomial distribution.

The **Poisson distribution** is used when it is desirable to determine the probability of obtaining *x* occurrences over an interval of time or space. The following assumptions are necessary for the Poisson distribution to be applicable.

1. The probability of an occurrence of the event is the same for any two intervals of equal length.

2. The occurrence or nonoccurrence of the event in any interval is independent of the occurrence or nonoccurrence of the event in any other interval.

A third discrete probability distribution, the **hypergeometric**, was introduced. Like the binomial, it is used to compute the probability of *x* successes in *n* trials. But, in contrast to the binomial, the probability of success changes from trial to trial.

## *Glossary*

**Random variable** A numerical description of the outcome of an experiment.

**Discrete random variable** A random variable that may assume either a finite number of values or an infinite sequence of values.

**Continuous random variable** A random variable that may assume any numerical value in an interval or collection of intervals.

**Probability distribution** A description of how the probabilities are distributed over the values of the random variable.

**Probability function** A function, denoted by $f(x)$, that provides the probability that *x* assumes a particular value for a discrete random variable.

**Discrete uniform probability distribution** A probability distribution for which each possible value of the random variable has the same probability.

**Expected value** A measure of the central location of a random variable, mean.

**Variance** A measure of the variability, or dispersion, of a random variable.

**Standard deviation** The positive square root of the variance.

**Binomial experiment** An experiment having the four properties stated in summary.

**Binomial probability distribution** A probability distribution showing the probability of *x* successes in *n* trials of a binomial experiment.

**Binomial probability function** The function used to compute binomial probabilities.

**Poisson probability distribution** A probability distribution showing the probability of *x* occurrences of an event over a specified interval of time or space.

**Poisson probability function** The function used to compute Poisson probabilities.

**Hypergeometric probability distribution** A probability distribution showing the probability of *x* successes in *n* trials from a population *N* with *r* successes and $N-r$ failures.

**Hypergeometric probability function** The function used to compute hypergeometric probabilities.

### *Key formulas*

Discrete Uniform Probability Function

$$f(x) = \frac{1}{n} \tag{3.1}$$

where $n$ = the number of values the random variable may assume

Expected Value (Mean) of a Discrete Random Variable

$$E(x) = \mu = \sum xf(x) \tag{3.2}$$

Variance of a Discrete Random Variable

$$Var(x) = \sigma^2 = \sum (x-\mu)^2 f(x) \tag{3.3}$$

Number of Experimental Outcomes Providing Exactly x Successes in $n$ Trials

$$\binom{n}{x} = \frac{n!}{x!(n-x)!} \tag{3.4}$$

Binomial Probability Function

$$f(x) = \binom{n}{x} p^x (1-p)^{(n-x)} \tag{3.5}$$

Expected Value (Mean) for the Binomial Distribution

$$E(x) = \mu = np \tag{3.6}$$

Variance for the Binomial Distribution

$$Var(x) = \sigma^2 = np(1-p) \tag{3.7}$$

Poisson Probability function

$$f(x) = \frac{\mu^x e^{-\mu}}{x!} \tag{3.8}$$

where $\mu$ – expected value (mean). For Poisson process $\mu = \sigma^2$.

Hypergeometric Probability Function

$$f(x) = \frac{C_x^r C_{n-x}^{N-r}}{C_n^N}, \quad for\ 0 \le x \le r \tag{3.9}$$

Expected Value for the Hypergeometric Distribution

$$E(x) = \mu = n\left(\frac{r}{N}\right) \tag{3.10}$$

Variance for the Hypergeometric Distribution

$$Var(x) = \sigma^2 = n\left(\frac{r}{N}\right)\left(1-\frac{r}{N}\right)\left(\frac{N-n}{N-1}\right) \tag{3.11}$$

# 3.2. Continuous Probability Distributions

## Summary

This lecture extended the discussion of probability distributions to the case of continuous random variables. The major conceptual difference between discrete and continuous probability distributions involves the method of computing probabilities. With discrete distributions, the probability function *f(x)* provides the probability that the random variable *x* assumes various values. With continuous distributions, the probability density function *f*(*x*) does not provide probability values directly. Instead, probabilities are given by areas under the curve or graph of the probability density function *f*(*x*). Because the area under the curve above a single point is zero, we observe that the probability of any particular value is zero for a continuous random variable.

Three continuous probability distributions – the uniform, normal, and exponential distributions – were treated in detail. The normal distribution is used widely in statistical inference and will be used extensively throughout the remainder of the text.

## Glossary

**Probability density function** A function used to compute probabilities for a continuous random variable. The area under the graph of a probability density function over an interval represents probability.

**Uniform probability distribution** A continuous probability distribution for which the probability that the random variable will assume a value in any interval is the same for each interval of equal length.

**Normal probability distribution** A continuous probability distribution. Its probability density function is bell shaped and determined by its mean $\mu$ and standard deviation $\sigma$.

**Standard normal probability distribution** A normal distribution with a mean of zero and a standard deviation of one.

**Continuity correction factor** A value of .5 that is added to or subtracted from a value of *x* when the continuous normal distribution is used to approximate the discrete binomial distribution.

**Exponential probability distribution** A continuous probability distribution that is useful in computing probabilities for the time between independent random events.

## Key formulas

Uniform Probability Density Function

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & for\ a \le x \le b \\ 0 \quad , & elsewhere \end{cases} \tag{3.12}$$

Expected Value and Variance of a Uniformly Distributed Variable $a \le x \le b$

$$E(x) = \mu = \frac{a+b}{2}, \qquad Var(x) = \sigma^2 = \frac{(b-a)^2}{12} \tag{3.13}$$

Normal Probability Density Function

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{3.14}$$

Converting to the Standard Normal Distribution

$$z = \frac{x-\mu}{\sigma} \tag{3.15}$$

Exponential Probability Density Function

$$f(x) = \frac{1}{\mu} e^{-\frac{x}{\mu}} \quad for\ x \ge 0, \mu > 0 \tag{3.16}$$

Exponential Distribution: Cumulative Probabilities

$$P(x \le x_0) = 1 - e^{-\frac{x_0}{\mu}} \tag{3.17}$$

## *Lecture supplementary material*

FIGURE 6.4 AREAS UNDER THE CURVE FOR ANY NORMAL DISTRIBUTION

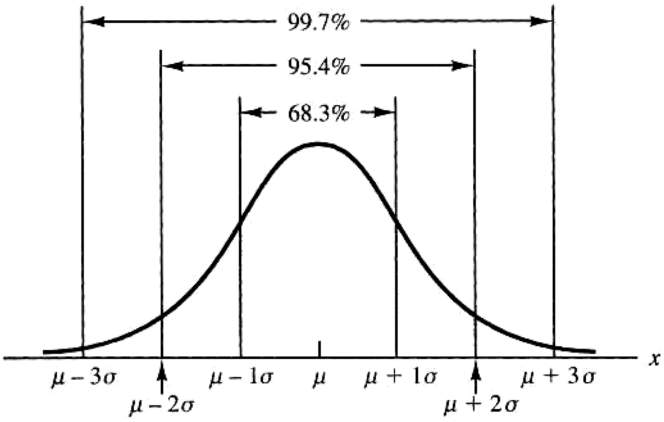


FIGURE 6.5 AREAS, OR PROBABILITIES, FOR THE STANDARD NORMAL DISTRIBUTION

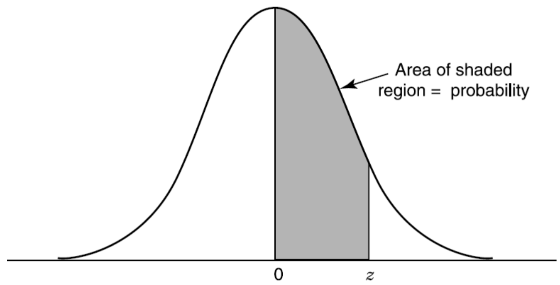


TABLE. Probability that a random variable has a value between 0 and z for the standard normal distribution (see Fig 6.5)

| z | .00 | .01 | .02 | .03 | .04 | .05 | .06 | .07 | .08 | .09 |
|---|---|---|---|---|---|---|---|---|---|---|
| .0 | .0000 | .0040 | .0080 | .0120 | .0160 | .0199 | .0239 | .0279 | .0319 | .0359 |
| .1 | .0398 | .0438 | .0478 | .0517 | .0557 | .0596 | .0636 | .0675 | .0714 | .0753 |
| .2 | .0793 | .0832 | .0871 | .0910 | .0948 | .0987 | .1026 | .1064 | .1103 | .1141 |
| .3 | .1179 | .1217 | .1255 | .1293 | .1331 | .1368 | .1406 | .1443 | .1480 | .1517 |
| .4 | .1554 | .1591 | .1628 | .1664 | .1700 | .1736 | .1772 | .1808 | .1844 | .1879 |
| .5 | .1915 | .1950 | .1985 | .2019 | .2054 | .2088 | .2123 | .2157 | .2190 | .2224 |
| .6 | .2257 | .2291 | .2324 | .2357 | .2389 | .2422 | .2454 | .2486 | .2517 | .2549 |
| .7 | .2580 | .2611 | .2642 | .2673 | .2704 | .2734 | .2764 | .2794 | .2823 | .2852 |
| .8 | .2881 | .2910 | .2939 | .2967 | .2995 | .3023 | .3051 | .3078 | .3106 | .3133 |
| .9 | .3159 | .3186 | .3212 | .3238 | .3264 | .3289 | .3315 | .3340 | .3365 | .3389 |
| 1.0 | .3413 | .3438 | .3461 | .3485 | .3508 | .3531 | .3554 | .3577 | .3599 | .3621 |
| 1.1 | .3643 | .3665 | .3686 | .3708 | .3729 | .3749 | .3770 | .3790 | .3810 | .3830 |
| 1.2 | .3849 | .3869 | .3888 | .3907 | .3925 | .3944 | .3962 | .3980 | .3997 | .4015 |
| 1.3 | .4032 | .4049 | .4066 | .4082 | .4099 | .4115 | .4131 | .4147 | .4162 | .4177 |
| 1.4 | .4192 | .4207 | .4222 | .4236 | .4251 | .4265 | .4279 | .4292 | .4306 | .4319 |
| 1.5 | .4332 | .4345 | .4357 | .4370 | .4382 | .4394 | .4406 | .4418 | .4429 | .4441 |
| 1.6 | .4452 | .4463 | .4474 | .4484 | .4495 | .4505 | .4515 | .4525 | .4535 | .4545 |
| 1.7 | .4554 | .4564 | .4573 | .4582 | .4591 | .4599 | .4608 | .4616 | .4625 | .4633 |
| 1.8 | .4641 | .4649 | .4656 | .4664 | .4671 | .4678 | .4686 | .4693 | .4699 | .4706 |
| 1.9 | .4713 | .4719 | .4726 | .4732 | .4738 | .4744 | .4750 | .4756 | .4761 | .4767 |
| 2.0 | .4772 | .4778 | .4783 | .4788 | .4793 | .4798 | .4803 | .4808 | .4812 | .4817 |
| 2.1 | .4821 | .4826 | .4830 | .4834 | .4838 | .4842 | .4846 | .4850 | .4854 | .4857 |
| 2.2 | .4861 | .4864 | .4868 | .4871 | .4875 | .4878 | .4881 | .4884 | .4887 | .4890 |
| 2.3 | .4893 | .4896 | .4898 | .4901 | .4904 | .4906 | .4909 | .4911 | .4913 | .4916 |
| 2.4 | .4918 | .4920 | .4922 | .4925 | .4927 | .4929 | .4931 | .4932 | .4934 | .4936 |
| 2.5 | .4938 | .4940 | .4941 | .4943 | .4945 | .4946 | .4948 | .4949 | .4951 | .4952 |
| 2.6 | .4953 | .4955 | .4956 | .4957 | .4959 | .4960 | .4961 | .4962 | .4963 | .4964 |
| 2.7 | .4965 | .4966 | .4967 | .4968 | .4969 | .4970 | .4971 | .4972 | .4973 | .4974 |
| 2.8 | .4974 | .4975 | .4976 | .4977 | .4977 | .4978 | .4979 | .4979 | .4980 | .4981 |
| 2.9 | .4981 | .4982 | .4982 | .4983 | .4984 | .4984 | .4985 | .4985 | .4986 | .4986 |
| 3.0 | .4987 | .4987 | .4987 | .4988 | .4988 | .4989 | .4989 | .4989 | .4990 | .4990 |