PROBABILITY DISTRIBUTION (JOINT PROBABILITIES DISTRIBUTION)

Probability and Statistics

6-1: Concept of Joint Probabilities

- Some random variables are not independent of each other , i.e., they tend to be related.

  \- Urban atmospheric ozone and airborne particulate matter tend to vary together.

  \- Urban vehicle speeds and fuel consumption rates tend to vary inversely.

6-1: Concept of Joint Probabilities

- The length (X) of a injection-molded part might not be independent of the width (*Y*) . Individual parts will vary due to random variation in materials and pressure.

- A joint probability distribution will describe the behavior of several random variables, say, *X* and *Y*. The graph of the distribution is 3-dimensional: x, y , and $f(x, y)$.

6-1: Concept of Joint Probabilities

Example 6-1: Signal Bars

You use your cell phone to check your airline reservation. The airline system requires that you speak the name of your departure city to the voice recognition system.

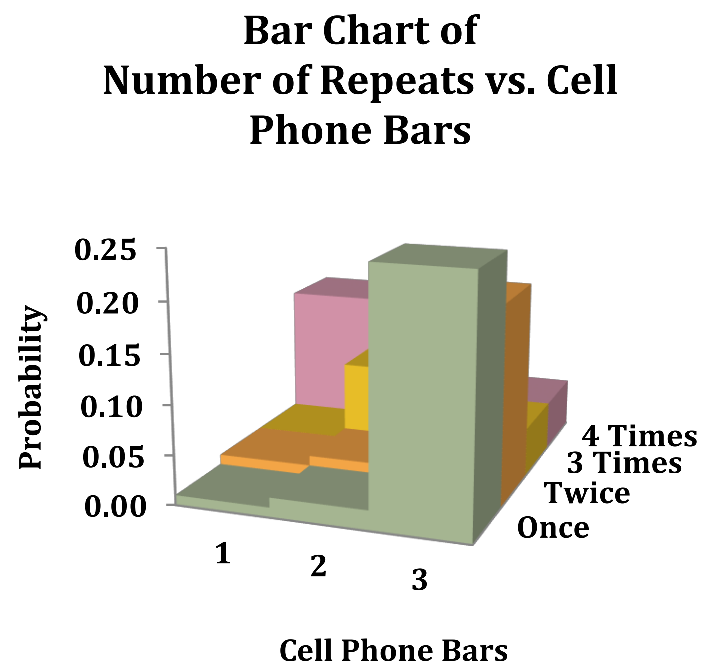
- Let Y denote the number of times that you have to state your departure city.
- Let X denote the number of bars of signal strength on you cell phone.

6-1: Concept of Joint Probabilities

Example 6-1: Signal Bars

y = number of times city name is stated	bars of signal strength		
	1	2	3
1	0.01	0.02	0.25
2	0.02	0.03	0.20
3	0.02	0.10	0.05
4	0.15	0.10	0.05

Figure 6-1 Joint probability distribution of X and Y. The table cells are the probabilities. Observe that more bars relate to less repeating.



6-1: Concept of Joint Probabilities

6-1.1 Joint Probability Mass Function Defined

The **joint probability mass function** of the discrete random variables X and Y denote as $f_{xy}(x, y)$, satisfies

(1) $f_{XY}(x, y) \geq 0$ All probability are non-negative

(2) $\sum_{x} \sum_{y} f_{XY}(x, y) = 1$ The sum of all probability is 1

(3) $f_{XY}(x, y) = P(X = x, Y = y)$

6-1: Concept of Joint Probabilities

6-1.2 Joint Probability Density Function Defined

The **joint probability density function** for the continuous random variables *X* and *Y*, denotes as $f_{XY}(x,y)$, satisfies the following properties:

$$(1)\ f_{XY}(x,y) \geq 0 \text{ for all } x,\ y$$

$$(2)\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)\,dxdy = 1$$

$$(3)\ P\big((X,Y) \subset R\big) = \iint_R f_{XY}(x,y)\,dxdy$$

6-1: Concept of Joint Probabilities

6-1.1 Joint Probability Density Function Defined

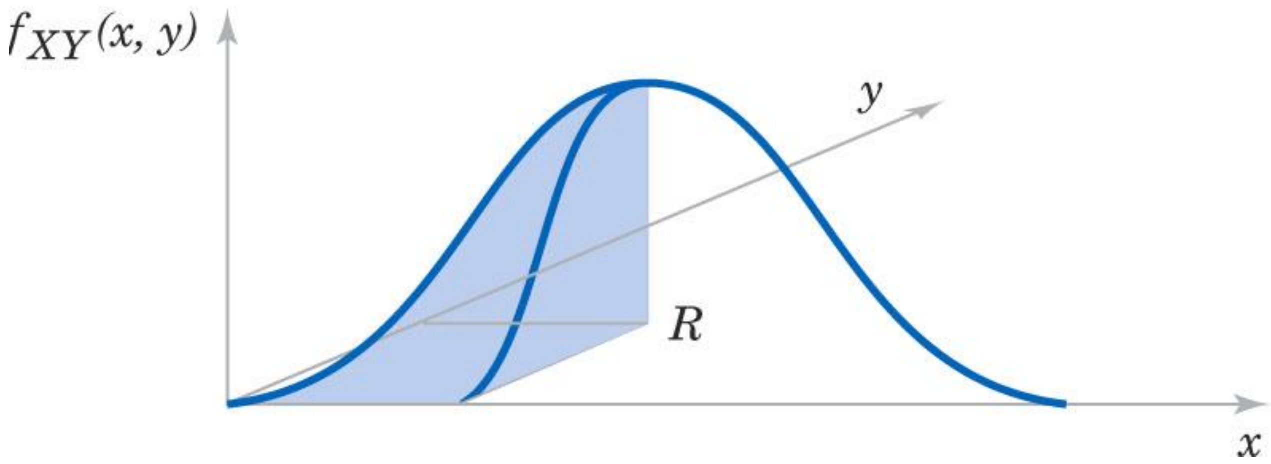


Figure 6-2 Joint probability density function for the random variables X and Y. Probability that (X, Y) is in the region R is determined by the **volume** of $f_{XY}(x,y)$ over the region R.

6-1: Concept of Joint Probabilities

6-1.2 Joint Probability Density Function Defined

The **joint probability density function** for the continuous random variables *X* and *Y*, denotes as $f_{XY}(x,y)$, satisfies the following properties:

$$(1)\ f_{XY}(x,y) \geq 0 \text{ for all } x,\ y$$

$$(2)\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)\,dxdy = 1$$

$$(3)\ P\big((X,Y) \subset R\big) = \iint_R f_{XY}(x,y)\,dxdy$$

6-1: Concept of Joint Probabilities

6-1.2 Joint Probability Density Function Defined

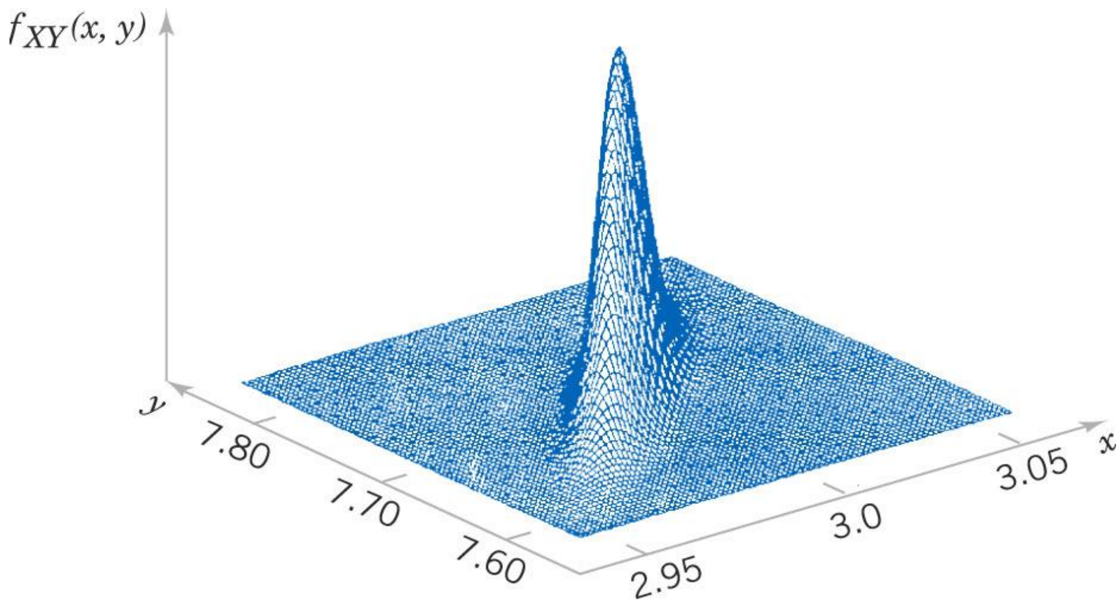


Figure 6-3 Joint probability density function for the continuous random variables X and Y of different dimensions of an injection-molded part. Note the asymmetric, narrow ridge shape of the PDF – indicating that small values in the X dimension are more likely to occur when small values in the Y dimension occur.

6-1: Concept of Joint Probabilities

Example 6-2: Server Access Time

Let the random variable X denote the time (msec's) until a computer server connects to your machine. Let Y denote the time until the server authorizes you as a valid user. X and Y measure the wait from a common starting point ($x < y$). The range of x and y are shown here.

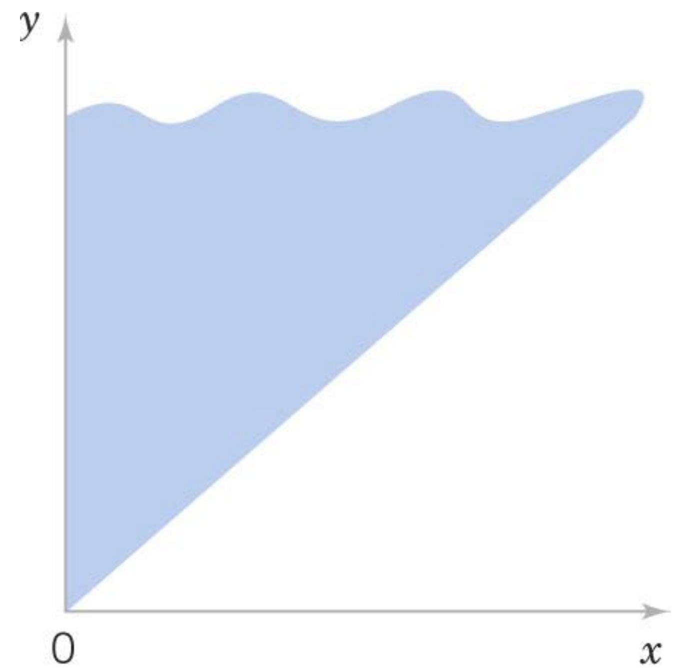


Figure 6-4 The joint probability density function of X and Y is nonzero over the shaded region where $x < y$.

6-1: Concept of Joint Probabilities

Example 6-2: Server Access Time

- The joint probability density function is:

$$f_{XY}(x,y) = ke^{-0.001x-0.002y} \text{ for } 0 < x < y < \infty \text{ and } k = 6\cdot 10^{-6}$$

- We verify that it integrates to 1 as follows:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)\,dx\,dy = \int_{0}^{\infty}\left(\int_{x}^{\infty} ke^{-0.001x-0.002y}\,dy\right)dx = k\int_{0}^{\infty}\left(\int_{x}^{\infty} e^{-0.002y}\,dy\right)e^{-0.001x}\,dx$$

$$= k\int_{0}^{\infty}\left(\frac{e^{-0.002x}}{0.002}\right)e^{-0.001x}\,dx = 0.003\int_{0}^{\infty} e^{-0.003x}\,dx$$

$$= 0.003\left(\frac{1}{0.003}\right) = 1$$

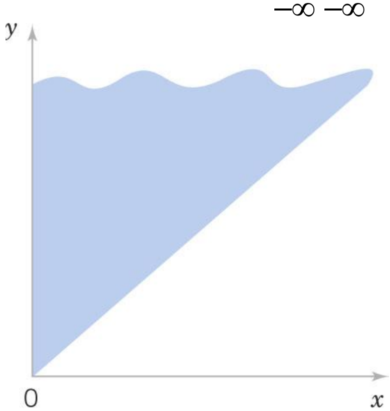


Figure 6-4 The joint PDF of X and Y is nonzero over the shaded region where $x < y$.

6-1: Concept of Joint Probabilities

Example 6-2: Server Access Time

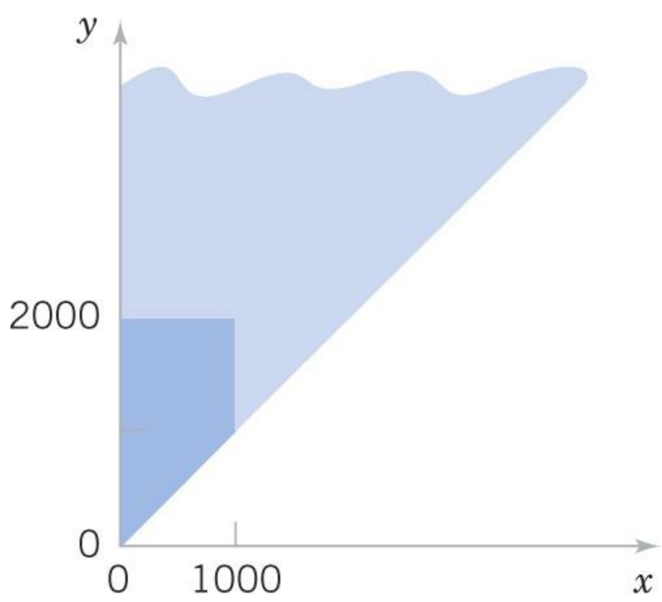


Figure 6-5 Region of integration for the probability that $X < 1000$ and $Y < 2000$ is darkly shaded.

Now calculate a probability:

$$P(X \le 1000, Y \le 2000) = \int_{-\infty}^{1000} \int_{x}^{2000} f_{XY}(x, y)\,dx\,dy$$

$$= k \int_{0}^{1000} \left(\int_{x}^{2000} e^{-0.002y}\,dy \right) e^{-0.001x}\,dx$$

$$= k \int_{0}^{1000} \left(\frac{e^{-0.002x} - e^{-4}}{0.002} \right) e^{-0.001x}\,dx$$

$$= 0.003 \int_{0}^{1000} e^{-0.003x} - e^{-4} e^{-0.001x}\,dx$$

$$= 0.003 \left[\left(\frac{1 - e^{-3}}{0.003} \right) - e^{-4} \left(\frac{1 - e^{-1}}{0.001} \right) \right]$$

$$= 0.003(316.738 - 11.578) = 0.915$$

6-2: Marginal Probability Distributions

6-2.1 Marginal Probability Distributions (discrete)

For a **discrete** joint PDF, there are marginal distributions for each random variable, formed by summing the joint PMF over the other variable.

$$f_X(x) = \sum_y f(xy) \qquad (6\text{-}1)$$

$$f_Y(y) = \sum_x f(xy) \qquad (6\text{-}2)$$

y = number of times city name is stated	x = number of bars of signal strength			
	1	2	3	$f(y)$ =
1	0.01	0.02	0.25	0.28
2	0.02	0.03	0.20	0.25
3	0.02	0.10	0.05	0.17
4	0.15	0.10	0.05	0.30
$f(x)$ =	0.20	0.25	0.55	1.00

Table 6-1 From the prior example, the joint PMF is shown in green while the two marginal PMFs are shown in lilac.

6-2: Marginal Probability Distributions

6-2.2 Marginal Probability Distributions (continuous)

- Rather than summing a discrete joint PMF, we integrate a continuous joint PDF.
- The marginal PDFs are used to make probability statements about one variable.
- If the joint probability density function of random variables X and Y is $f_{XY}(x,y)$, the marginal probability density functions of X and Y are:

$$f_X(x) = \int_y f_{XY}(x,y)\,dy \qquad (6\text{-}3)$$

$$f_Y(y) = \int_x f_{XY}(x,y)\,dx \qquad (6\text{-}4)$$

6-2: Marginal Probability Distributions

Example 6-2: Server Access Time

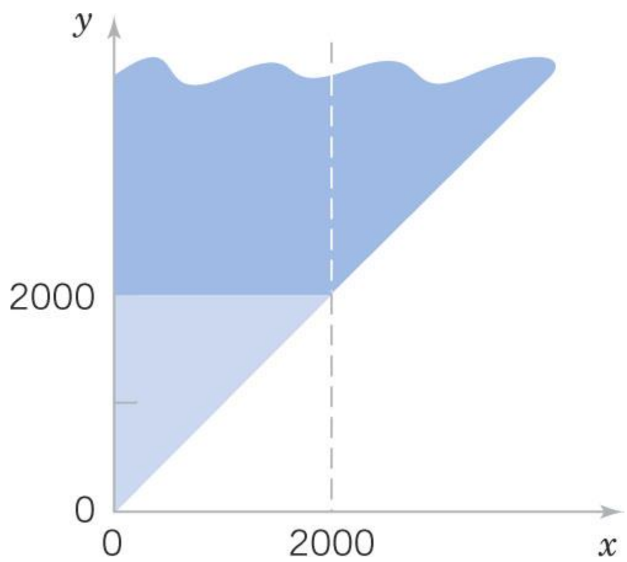


Figure 6-6

For the random variables times in Example 6-2, find the probability that Y exceeds 2000.

Integrate the joint PDF directly using the picture to determine the limits.

$$P(Y \geq 2000) = \int_{0}^{2000} \left(\int_{2000}^{\infty} f_{XY}(x,y)\,dy \right) dx + \int_{2000}^{\infty} \left(\int_{x}^{\infty} f_{XY}(x,y)\,dy \right) dx$$

Dark region = left dark region + right dark region

6-2: Marginal Probability Distributions

Example 6-2: Server Access Time

Alternatively, find the marginal PDF and then integrate that to find the desired probability.

$$f_Y(y) = \int_0^y ke^{-0.001x-0.002y}$$

$$= ke^{-0.002y} \int_0^y e^{=0.001x} dx$$

$$= ke^{-0.002y} \left(\left. \frac{e^{-0.001x}}{-0.001} \right|_0^y \right)$$

$$= ke^{-0.002y} \left(\frac{1-e^{-0.001y}}{0.001} \right)$$

$$= 6 \cdot 10^{-3} e^{-0.002y} \left(1-e^{-0.001y} \right) \text{ for } y > 0$$

$$P(Y \geq 2000) = \int_{2000}^{\infty} f_Y(y) dy$$

$$= 6 \cdot 10^{-3} \int_{2000}^{\infty} e^{-0.002y} \left(1-e^{-0.001y} \right) dy$$

$$= 6 \cdot 10^{-3} \left[\left(\left. \frac{e^{-0.002y}}{-0.002} \right|_{2000}^{\infty} \right) - \left(\left. \frac{e^{-0.003y}}{-0.003} \right|_{2000}^{\infty} \right) \right]$$

$$= 6 \cdot 10^{-3} \left[\frac{e^{-4}}{0.002} - \frac{e^{-6}}{0.003} \right] = 0.05$$

6-2: Marginal Probability Distributions

6-2.3 Mean & Variance of a Marginal Distribution

Means E(X) and E(Y) are calculated from the discrete and continuous marginal distributions.

	Discrete	Continuous	
$E(X)=\sum_R x\cdot f_X(x)$		$=\int_R x\cdot f_X(x)\,dx=\mu_X$	(6-5)
$E(Y)=\sum_R y\cdot f_Y(y)$		$=\int_R y\cdot f_Y(y)\,dy=\mu_Y$	(6-6)
$V(X)=\sum_R x^2\cdot f_X(x)-\mu_X^2$		$=\int_R x^2\cdot f_X(x)\,dx-\mu_X^2$	(6-7)
$V(Y)=\sum_R y^2\cdot f_Y(y)-\mu_Y^2$		$=\int_R y^2\cdot f_Y(y)\,dy-\mu_Y^2$	(6-8)

6-2: Marginal Probability Distributions

6-2.3 Mean & Variance of a Marginal Distribution

Example 6-3: Signal Bars (Cont.)

y = number of times city name is stated	x = number of bars of signal strength					
	1	2	3	$f(y)$ =	$y*f(y)$ =	$y^2*f(y)$ =
1	0.01	0.02	0.25	0.28	0.28	0.28
2	0.02	0.03	0.20	0.25	0.50	1.00
3	0.02	0.10	0.05	0.17	0.51	1.53
4	0.15	0.10	0.05	0.30	1.20	4.80
$f(x)$ =	0.20	0.25	0.55	1.00	2.49	7.61
$x*f(x)$ =	0.20	0.50	1.65	2.35		
$x^2*f(x)$ =	0.20	1.00	4.95	6.15		

$E(X) = 2.35 \quad V(X) = 6.15 - 2.35^2 = 6.15 - 5.52 = 0.6275$

$E(Y) = 2.49 \quad V(Y) = 7.61 - 2.49^2 = 7.61 - 16.20 = 1.4099$

6-3: Conditional Probability Density Function Defined

6-3.1 Conditional Probability Distributions

Given random variables X and Y with joint probability density function $f_{XY}(x, y)$, the conditional probability densiy function of Y given X=x is

$$f_{Y|x}(y) = \frac{f_{XY}(x, y)}{f_X(x)} \quad \text{for} \quad f_X(x) > 0 \qquad (6\text{-}9)$$

which satifies the following properties:

(1) $f_{Y|x}(y) \geq 0$

(2) $\int f_{Y|x}(y)\,dy = 1$ or $\sum f_{Y|x}(y) = 1$

(3) $P(Y \subset B | X = x) = \begin{cases} \int_B f_{Y|x}(y)\,dy \\ \sum\limits_B f_{Y|x}(y) \end{cases}$, for any set B in the range of Y

6-3: Conditional Probability Density Function Defined

6-3.1 Conditional Probability Distributions

Example 6-4: Signal Bars (Cont. - Discrete)

Recall that $P(B|A) = \dfrac{P(A \cap B)}{P(A)}$

From **Example 6-3**

P(Y=1|X=3) = 0.25/0.55 = 0.455

P(Y=2|X=3) = 0.20/0.55 = 0.364

P(Y=3|X=3) = 0.05/0.55 = 0.091

P(Y=4|X=3) = 0.05/0.55 = 0.091

Sum = 1.001

y = number of times city name is stated	x = number of bars of signal strength			
	1	2	3	$f(y)$ =
1	0.01	0.02	0.25	0.28
2	0.02	0.03	0.20	0.25
3	0.02	0.10	0.05	0.17
4	0.15	0.10	0.05	0.30
$f(x)$ =	0.20	0.25	0.55	1.00

Note that there are 12 probabilities conditional on *X*, and 12 more probabilities conditional upon *Y*.

6-3: Conditional Probability Density Function Defined

6-3.1 Conditional Probability Distributions

Example 6-4: Signal Bars (Cont. - Discrete)

From **Example 6-3:** Conditional discrete PMFs can be shown as tables.

y = number of times city name is stated	x = number of bars of signal strength				f(x|y) for y =			Sum of f(x|y) =
	1	2	3	$f(y)$ =	1	2	3	
1	0.01	0.02	0.25	0.28	0.036	0.071	0.893	1.000
2	0.02	0.03	0.20	0.25	0.080	0.120	0.800	1.000
3	0.02	0.10	0.05	0.17	0.118	0.588	0.294	1.000
4	0.15	0.10	0.05	0.30	0.500	0.333	0.167	1.000
$f(x)$ =	0.20	0.25	0.55					
1	0.050	0.080	0.455					
2	0.100	0.120	0.364					
3	0.100	0.400	0.091					
4	0.750	0.400	0.091					
Sum of f(y|x) =	1.000	1.000	1.000					

6-3: Conditional Probability Density Function Defined

6-3.1 Conditional Probability Distributions

Example 6-5: Server Access Time (Cont. - Continuous)

Determine the conditional PDF for Y given X=x.

$$f_X(x) = \int_x^{\infty} k \cdot e^{-0.001x-0.002y} dy$$

$$= ke^{-0.001x}\left(\left. \frac{e^{-0.002y}}{0.002} \right|_x^{\infty} \right)$$

$$= ke^{-0.001x}\left(\frac{e^{-0.002}}{0.002} \right)$$

$$= 0.003e^{-0.003x} \quad \text{for} \quad x > 0$$

$$f_{Y|x}(y) = \frac{f_{XY}(x,y)}{f_X((x))} = \frac{ke^{-0.001x-0.002y}}{0.003e^{-0.003x}}$$

$$= 0.002e^{0.002x-0.002y} \quad \text{for } 0 < x < y < \infty$$

6-3: Conditional Probability Density Function Defined

6-3.1 Conditional Probability Distributions

Example 6-5: Server Access Time (Cont. - Continuous)

Now find the probability that Y exceeds 2000 given that X=1500:

$$P(Y > 2000 | X = 1500)$$

$$= \int_{2000}^{\infty} f_{Y|1500}(y)\,dy$$

$$= \int_{2000}^{\infty} 0.002e^{0.002(1500)-0.002y}$$

$$= 0.002e^{3}\left(\left.\frac{e^{-0.002y}}{-0.002}\right|_{2000}^{\infty}\right)$$

$$= 0.002e^{3}\left(\frac{e^{-4}}{0.002}\right) = e^{-1} = 0.368$$

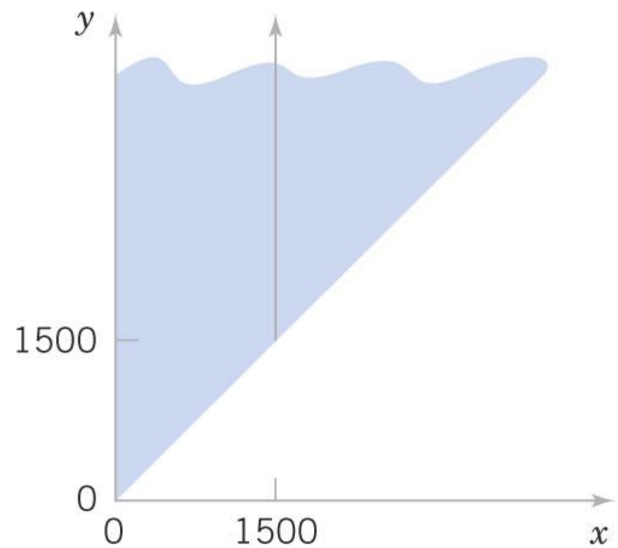


Figure 6-7 **Again,** the conditional PDF is nonzero on the solid line in the shaded region.

6-3: Conditional Probability Density Function Defined

6-3.2 Mean & Variance of Conditional Random Variables

- The conditional mean of Y given $X = x$, denoted as $E(Y|x)$ or $\mu_{Y|x}$ is:

$$E(Y|x) = \int_y y \cdot f_{Y|x}(y)\,dy \qquad (6\text{-}10)$$

- The conditional variance of Y given $X = x$, denoted as $V(Y|x)$ or $\sigma^2_{Y|x}$ is:

$$V(Y|x) = \int_y \left(y - \mu_{Y|x}\right)^2 \cdot f_{Y|x}(y)\,dy = \int_y y^2 \cdot f_{Y|x}(y) - \mu^2_{Y|x} \qquad (6\text{-}11)$$

6-3.2 Mean & Variance of Conditional Random Variables

Example 6-6: Server Access Time (Cont. - Continuous)

What is the conditional mean for Y given that $x = 1500$?

Integrate by parts.

$$E(Y|X=1500) = \int_{1500}^{\infty} y \cdot 0.002e^{0.002(1500)-0.002y}\,dy = 0.002e^{3}\int_{1500}^{\infty} y \cdot e^{-0.002y}\,dy$$

$$= 0.002e^{3}\left[y\frac{e^{-0.002y}}{-0.002}\Bigg|_{1500}^{\infty} - \int_{1500}^{\infty}\left(\frac{e^{-0.002y}}{-0.002}\right)dy \right]$$

$$= 0.002e^{3}\left[\frac{1500}{0.002}e^{-3} - \left(\frac{e^{-0.002y}}{(0.002)(0.002)}\Bigg|_{1500}^{\infty} \right) \right]$$

$$= 0.002e^{3}\left[\frac{1500}{0.002}e^{-3} + \frac{e^{-3}}{(0.002)(0.002)} \right]$$

$$= 0.002e^{3}\left[\frac{e^{-3}}{0.002}(2000) \right] = 2000$$

If the connect time is 1500 ms, then the expected time to be authorized is 2000 ms.

6-3: Conditional Probability Density Function Defined

6-3.2 Mean & Variance of Conditional Random Variables

Example 6-7: Signal Bars (Cont. - Discrete)

For the discrete random variables in Exercise 6-1, what is the conditional mean of Y given X=1?

y = number of times city name is stated	x = number of bars of signal strength				
	1	2	3	$f(y)$ =	
1	0.01	0.02	0.25	0.28	
2	0.02	0.03	0.20	0.25	
3	0.02	0.10	0.05	0.17	
4	0.15	0.10	0.05	0.30	
$f(x)$ =	0.20	0.25	0.55	y*f(y|x=1)	y^2*f(y|x=1)
1	0.050	0.080	0.455	0.05	0.05
2	0.100	0.120	0.364	0.20	0.40
3	0.100	0.400	0.091	0.30	0.90
4	0.750	0.400	0.091	3.00	12.00
Sum of f(y|x) =	1.000	1.000	1.000	3.55	13.35
					12.6025
					0.7475

The mean number of attempts given one bar is 3.55 with variance of 0.7475.

6-4: More Than One Random Variable

6-4.1 Independence of Joint Random Variable

Random variable independence means that knowledge of the values of X does not change any of the probabilities associated with the values of Y.

- *X* and *Y* vary independently.
- Dependence implies that the values of *X* are influenced by the values of *Y*.
- Do you think that a person's height and weight are independent?

The random variables $X_1, X_2, \ldots, X_n$ are **independent** if

$$P(X_1 \in E_1, X_2 \in E_2, \ldots, X_n \in E_n) = P(X_1 \in E_1)P(X_2 \in E_2) \cdots P(X_n \in E_n)$$

for *any* sets $E_1, E_2, \ldots, E_n$. (6-12)

6-4: More Than One Random Variable

Example 6-7: Optical Drive Diameter (Independence)

The probability that a diameter meets specifications (0.2485, 0.2515) was determined. What is the probability that 10 diameters all meet specifications, assuming that the diameters are independent?

Solution. Denote the diameter of the first shaft as X_1, the diameter of the second shaft as X_2, and so forth, so that the diameter of the tenth is denote as X_{10}. The probability that all shafts meet specifications can be written as

$$P\left(0.2485 < X_1 < 0.2515, 0.2485 < X_2 < 0.2515, \ldots, 0.2485 < X_{10} < 0.2515\right)$$

In this example, the only set of integers is

$$E_1 = \left(0.2485, 0.2515\right)$$

With respect to the notation used in the definition of independence, $E_1 = E_2 = \ldots = E_{10}$

6-4: More Than One Random Variable

Example 6-8: Relative Frequency (Independence)

Recall the relative frequency interpretation of probability. The proportion of times that shaft 1 is expected to meet the specifications is 0.919 [$P(0.2485 < x_1 < 0.2515) = 0.919$], the proportion of times that shaft 2 is expected to meet the specifications is 0.919, and so forth. If the random variables are independent , the proportion od times in which we measure 10 shaft that we expect all to meet the specifications is

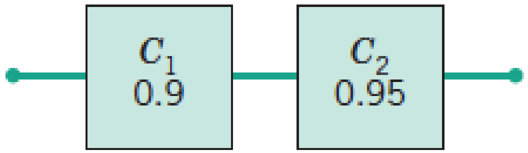
$$= P(0.2485 < X_1 < 0.2515) \times P(0.2485 < X_2 < 0.2515) \times \ldots \times P(0.2485 < X_{10} < 0.2515)$$

$$= 0.919^{10} = 0.430$$

6-4: More Than One Random Variable

Example 6-9: Series System (Independence)

The system show here operates only if there is a path of functional components from left to right. The probability that each component functions is show in the diagram. Assume that the components function or fail independently. What is the probability that the system operates?



Solution. Let C_1 and C_2 denote the events that components 1 and 2 are functional, respectively. For the system to operate, both components must be functional. The probability that the system operates is

$$P(C_1, C_2) = P(C_1)P(C_2) = (0.9)(0.95) = 0.855$$

Note that the probability that the system operates is smaller than the probability that any component operates. This system fails whenever *any* component fails. A system of this type is call s **series system**

6-4: More Than One Random Variable

Example 6-10: Plastic Molding (Independence)

In a plastic molding operation, each part is classified as to whether it conforms to color and length specifications.

$$X = \begin{cases} 1 \text{ if the part conforms to color specs} \\ 0 \text{ otherwise} \end{cases}$$

$$Y = \begin{cases} 1 \text{ if the part conforms to length specs} \\ 0 \text{ otherwise} \end{cases}$$

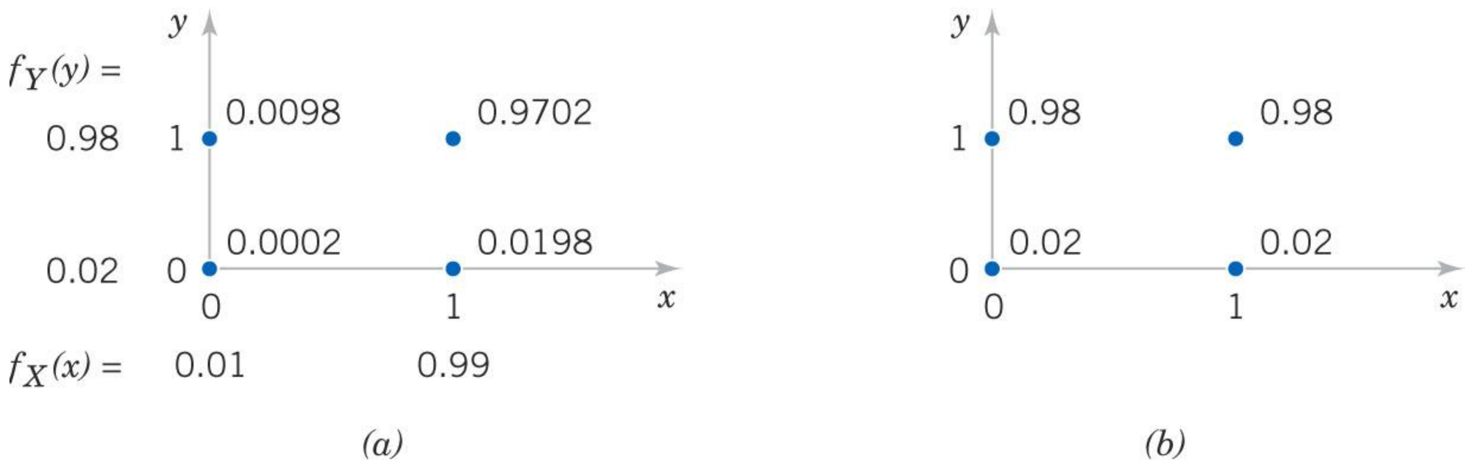


Figure 6-8(a) shows marginal & joint probabilities, $f_{XY}(x, y) = f_X(x) * f_Y(y)$

Figure 6-8(b) show the conditional probabilities, $f_{Y|x}(y) = f_Y(y)$

6-6: Covariance & Correlation

6-6.1 Covariance Defined

- Covariance is a measure of the relationship between two random variables
- First, we need to describe the expected value of a function of two random variables. Let *h*(*X*, *Y*) denote the function of interest.

$$E\left[h(X,Y)\right]=\begin{cases}\sum\sum h(x,y)\cdot f_{XY}(x,y)\text{ for } X,Y \text{ discrete}\\ \iint h(x,y)\cdot f_{XY}(x,y)\,dxdy \text{ for } X,Y \text{ continuous}\end{cases} \qquad (6\text{-}17)$$

6-6: Covariance & Correlation

6-6.1 Covariance Defined

The covariance between the random variables X and Y, denoted as $\text{cov}(X,Y)$ or σ_{XY} is

$$\sigma_{XY} = E\left[\left(X-\mu_X\right)\left(Y-\mu_Y\right)\right] = E\left(XY\right) - \mu_X\mu_Y \qquad (6\text{-}18)$$

The units of σ_{XY} are units of X times units of Y.

For example, if the units of X are feet
and the units of Y are pounds,
the units of the covariance are foot-pounds.

Unlike the range of variance, $-\infty < \sigma_{XY} < \infty$.

6-6: Covariance & Correlation

6-6.1 Covariance Defined

Example 6-17: E(Function of 2 Random Variables)

Task: Calculate $E[(X-\mu_X)(Y-\mu_Y)] = \text{cov}(X,Y)$

	x	y	f(x, y)	x-μ_X	y-μ_Y	Prod
Joint	1	1	0.1	-1.4	-1.0	0.14
	1	2	0.2	-1.4	0.0	0.00
	3	1	0.2	0.6	-1.0	-0.12
	3	2	0.2	0.6	0.0	0.00
	3	3	0.3	0.6	1.0	0.18
Marginal	1		0.3	covariance =		0.20
	3		0.7			
		1	0.3			
		2	0.4			
		3	0.3			
Mean		μ_X =	2.4	$= (1(0.3) + 3(0.7))$		
		μ_Y =	2.0	$= 1(0.3) + 2(0.4) + 3(0.3)$		

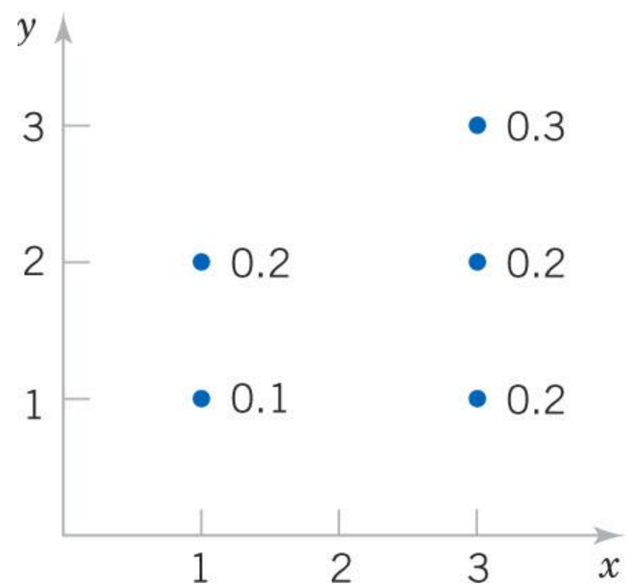


Figure 6-9 Discrete joint distribution of X and Y.

6-6: Covariance & Correlation

6-6.2 Covariance and Scatter Patterns

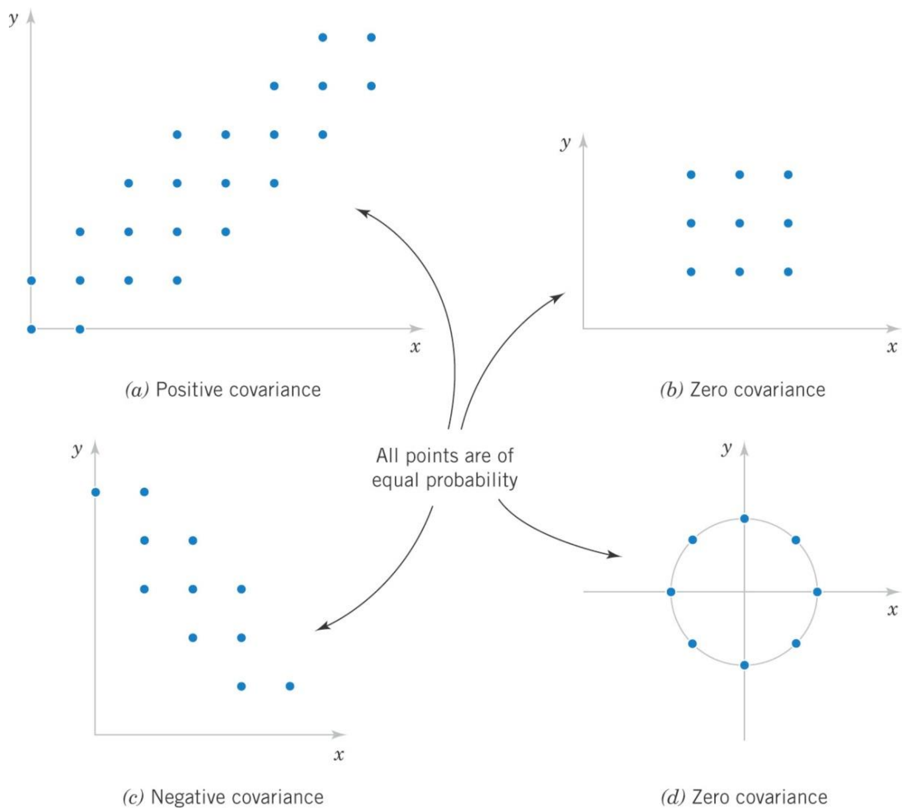


Figure 6-10 Joint probability distributions and the sign of cov(X, Y). Note that covariance is a measure of linear relationship. Variables with non-zero covariance are **correlated.**

6-6: Covariance & Correlation

6-6.1 Covariance Defined

Example 6-18: Signal Bars (Cont.)

y = number of times city name is stated	x = number of bars of signal strength		
	1	2	3
1	0.01	0.02	0.25
2	0.02	0.03	0.20
3	0.02	0.10	0.05
4	0.15	0.10	0.05

The probability distribution of **Example 6-1** is shown.

By inspection, note that the larger probabilities occur as *X* and *Y* move in opposite directions. This indicates a negative covariance.

6-6: Covariance & Correlation

6-6.2 Correlation (ρ = rho)

The correlation between random variables X and Y, denoted as ρ_{XY}, is

$$\rho_{XY} = \frac{\operatorname{cov}(X,Y)}{\sqrt{V(X)\cdot V(Y)}} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \qquad (6\text{-}19)$$

Since $\sigma_X > 0$ and $\sigma_Y > 0$,

ρ_{XY} and $\operatorname{cov}(X,Y)$ have the same sign.

We say that ρ_{XY} is normalized, so $-1 \le \rho_{XY} \le 1$

Note that ρ_{XY} is dimensionless.

Variables with non-zero correlation are correlated.

6-6: Covariance & Correlation

6-6.2 Correlation (ρ = rho)

Example 6-19:

Determine the covariance and correlation.

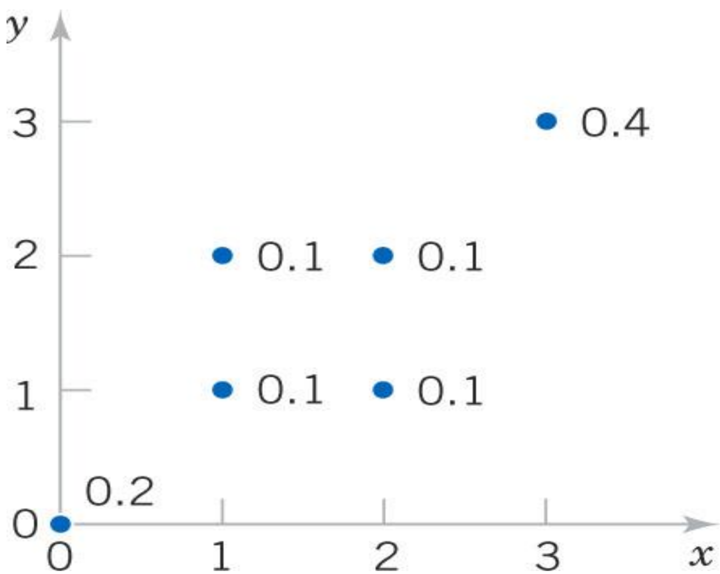


Figure 6-11 Discrete joint distribution, $f(x, y)$.

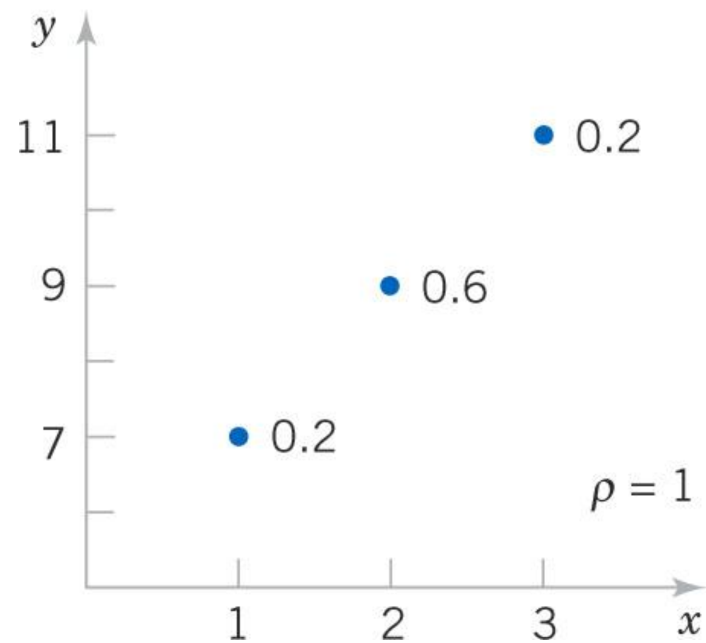
	x	y	f(x, y)	x-μ_X	y-μ_Y	Prod
Joint	0	0	0.2	-1.8	-1.2	0.42
	1	1	0.1	-0.8	-0.2	0.01
	1	2	0.1	-0.8	0.8	-0.07
	2	1	0.1	0.2	-0.2	0.00
	2	2	0.1	0.2	0.8	0.02
	3	3	0.4	1.2	1.8	0.88
Marginal	0		0.2		covariance =	1.260
	1		0.2		correlation =	0.926
	2		0.2			
	3		0.4	Note the strong positive correlation.		
		0	0.2			
		1	0.2			
		2	0.2			
		3	0.4			
Mean		μ_X =	1.8			
		μ_Y =	1.8			
StDev		σ_X =	1.1662			
		σ_Y =	1.1662			

6-6: Covariance & Correlation

6-6.2 Correlation (ρ = rho)

Example 6-20:

Determine the covariance and correlation.



	x	y	f(x, y)	x*y*f	
Joint	1	7	0.2	1.4	
	2	9	0.6	10.8	
	3	11	0.2	6.6	
Marginals	1		0.2	18.8	= E(XY)
	2		0.6	0.80	= $\mathrm{cov}(X,Y)$
	3		0.2	1.00	= ρ_{XY}
		7	0.2		
		9	0.6		
		11	0.2		
Mean		μ_X =	2		
		μ_Y =	9.0		
StDev		σ_X =	0.632		
		σ_Y =	1.265		

Figure 6-12 Discrete joint distribution. Steepness of line connecting points is immaterial.

6-4: More Than ~~One~~ two Random Variable

6-4.1 Independence of Joint Random Variable

Properties of Independence

For random variables X and Y, if any one of the following properties is true, the others are also true. Then X and Y are **independent.**

$$(1)\ f_{XY}(x,y)=f_X(x)\cdot f_Y(y)$$

$$(2)\ f_{Y|x}(y)=f_Y(y)\ \text{for all x and y with}\ f_X(x)>0$$

$$(3)\ f_{X|y}(y)=f_X(x)\ \text{for all x and y with}\ f_Y(y)>0$$

$$(4)\ \mathrm{P}(X\subset A, Y\subset B)=P(X\subset A)\cdot P(Y\subset B)\ \text{for any}$$

sets A and B in the range of X and Y, respectively.

6-4: More Than One Random Variable

6-4.1 Independence of Joint Random Variable

Rectangular Range for (*X*, *Y*)

- A rectangular range for *X* and *Y* is a necessary, but not sufficient, condition for the independence of the variables.
- If the range of X and Y is not rectangular, then the range of one variable is limited by the value of the other variable.
- If the range of X and Y is rectangular, then one of the properties in previous page must be demonstrated to prove independence.

6-4: More Than One Random Variable

6-4.1 Independence of Joint Random Variable

Example 6-11: Server Access Time (Cont. - Independence)

- Suppose the Example 6-5 is modified such that the joint PDF is:

$$f_{XY}(x,y) = 2\cdot 10^{-6} e^{=-0.001x-0.002y} \text{ for } x \ge 0 \text{ and } y \ge 0.$$

- Are X and Y independent? Is the product of the marginal PDFs equal the joint PDF? Yes by inspection.

$$f_X(x) = \int_0^\infty 2\cdot 10^{-6} e^{-0.001x-0.002y}\,dy \qquad f_Y(y) = \int_0^\infty 2\cdot 10^{-6} e^{-0.001x-0.002y}\,dx$$

$$= 0.001e^{-0.001x} \text{ for } x \ge 0 \qquad = 0.002e^{-0.002y} \text{ for y} \ge 0$$

- Find this probability: $P(X > 1000, Y < 1000) = P(X > 1000)\cdot P(Y < 1000)$

$$= e^{-1}\cdot\left(1-e^{-2}\right) = 0.318$$

6-4: More Than One Random Variable

6-4.1 Independence of Joint Random Variable

Example 6-12: Machined Dimensions

Let the random variables X and Y denote the lengths of 2 dimensions of a machined part. Assume that X and Y are independent and normally distributed. Find the desired probability.

	Normal Random Variables	
	X	Y
Mean	10.5	3.2
Variance	0.0025	0.0036

$$P(10.4 < X < 10.6,\ 3.15 < Y < 3.25) = P(10.4 < X < 10.6) \cdot P(3.15 < Y < 3.25)$$

$$= P\left(\frac{10.4-10.5}{0.05} < Z < \frac{10.6-10.5}{0.05}\right) \cdot P\left(\frac{3.15-3.2}{0.06} < Z < \frac{3.25-3.2}{0.06}\right)$$

$$= P(-2 < Z < 2) \cdot P(-0.833 < Z < 0.833) = 0.568$$

6-4: More Than One Random Variable

6-4.2 More Than Two Random Variables

- Many dimensions of a machined part are routinely measured during production. Let the random variables X_1, X_2, X_3 and X_4 denote the lengths of four dimensions of a part.

- What we have learned about joint, marginal and conditional PDFs in two variables extends to many (p) random variables.

6-4: More Than One Random Variable

6-4.2 More Than Two Random Variables

Joint Probability Density Function Redefined

The **joint probability density function** for the continuous random variables X_1, X_2, X_3, …X_p, denoted as satisfies the following properties:

$$(1)\ f_{X_1X_2\ldots X_p}\left(x_1,x_2,\ldots,x_p\right)\geq 0$$

$$(2)\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} f_{X_1X_2\ldots X_p}\left(x_1,x_2,\ldots,x_p\right)dx_1dx_2\ldots dx_p=1$$

(3) For any region B of p-dimensional space,

$$P\left(\left(X_1,X_2\ldots X_p\right)\subset B\right)=\iint_B\ldots\int f_{X_1X_2\ldots X_p}\left(x_1,x_2,\ldots,x_p\right)dx_1dx_2\ldots dx_p$$

6-4: More Than One Random Variable

6-4.2 More Than Two Random Variables

Example 6-13: Component Lifetimes

In an electronic assembly, let X_1, X_2, X_3, X_4 denote the lifetimes of 4 components in hours. The joint PDF is:

$$f_{X_1X_2X_3X_4}(x_1, x_2, x_3, x_4) = 9 \cdot 10^{-12} e^{-0.001x_1 - 0.002x_2 - 0.0015x_3 = 0.003x_4} \text{ for } x_i \geq 0$$

What is the probability that the device operates more than 1000 hours?

The joint PDF is a product of exponential PDFs.

$P(X_1 > 1000, X_2 > 1000, X_3 > 1000, X_4 > 1000)$

$= \quad e^{-1-2-1.5-3} = e^{-7.5} = 0.00055$

6-5: More Than Two Random Variables

6-5.1 Marginal Probability Density Function

If the joint probability density function of continuous random variables

$X_1, X_2, \ldots X_p$ is $f_{X_1 X_2 \ldots X_p}\left(x_1, x_2, \ldots x_p\right)$,

the marginal probability density function of X_i is

$$f_{X_i}\left(x_i\right) = \int\int \ldots \int f_{X_1 X_2 \ldots X_p}\left(x_1, x_2, \ldots x_p\right) dx_1 dx_2 \ldots dx_{i-1} dx_{i+1} \ldots dx_p \qquad (6\text{-}13)$$

where the integral is over all points in the range of $X_1, X_2, \ldots X_p$

for which $X_i = x_i$. (don't integrate out x_i)

6-5: More Than Two Random Variables

6-5.2 Mean & Variance of a Joint PDF

The mean and variance of X_i can be determined from either the marginal PDF, or the joint PDF as follows:

$$E(X_i) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_i \cdot f_{X_1 X_2 \ldots X_p}(x_1, x_2, \ldots x_p)\, dx_1 dx_2 \ldots dx_p$$

and (6-14)

$$V(X_i) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(x_i - \mu_{X_i}\right)^2 \cdot f_{X_1 X_2 \ldots X_p}(x_1, x_2, \ldots x_p)\, dx_1 dx_2 \ldots dx_p$$

6-5: More Than Two Random Variables

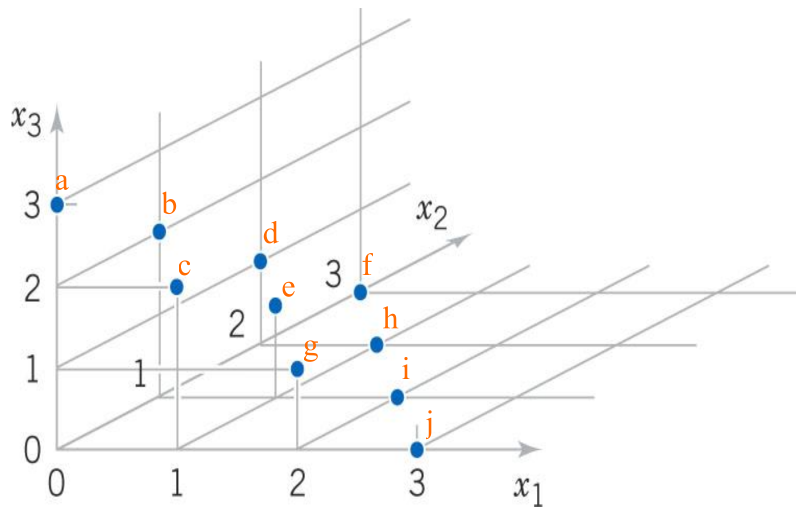
6-5.2 Mean & Variance of a Joint PDF

Example 6-15:

- There are 10 points in this discrete joint PDF.
- Note that

$$x_1+x_2+x_3 = 3$$

- List the marginal PDF of X_2



a+c+g+j $P(X_2 = 0) =$	$P_{XXX}(0,0,3) +$	$P_{XXX}(1,0,2) +$	$P_{XXX}(2,0,1) +$	$P_{XXX}(3,0,0)$
$P(X_2 = 1) =$	$P_{XXX}(0,1,2) +$	$P_{XXX}(1,1,1) +$	$P_{XXX}(2,1,0)$	
c+h $P(X_2 = 2) =$	$P_{XXX}(0,2,1) +$	$P_{XXX}(1,2,0)$		
$P(X_2 = 3) =$	$P_{XXX}(0,3,0)$		Note the index pattern	

6-5: More Than Two Random Variables

6-5.3 Reduced Dimensionality

If the joint probability density function of continuous random variables

$X_1, X_2, \ldots X_p$ is $f_{X_1 X_2 \ldots X_p}\left(x_1, x_2, \ldots x_p\right)$,

then the probability density function of $X_1, X_2, \ldots X_k$, $k < p$ is

$$f_{X_1 X_2 \ldots X_k}\left(x_1, x_2, \ldots, x_k\right) = \int\int\ldots\int f_{X_1 X_2 \ldots X_p}\left(x_1, x_2, \ldots x_p\right) dx_{k+1}\, dx_{k+2} \ldots dx_p \qquad (6\text{-}15)$$

where the integral is over all points in the

range of $X_1, X_2, \ldots X_p$ for which $X_i = x_i$ for $i = 1$ through k.

(integrate out p-k variables)

6-5: More Than Two Random Variables

6-5.4 Conditional Probability Distributions

- Conditional probability distributions can be developed for multiple random variables by extension of the ideas used for two random variables.
- Suppose $p = 5$ and we wish to find the distribution conditional on X_4 and X_5.

$$f_{X_1X_2X_3|X_4X_5}\left(x_1, x_2, x_3\right) = \frac{f_{X_1X_2X_3X_4X_5}\left(x_1, x_2, x_3, x_4, x_5\right)}{f_{X_4X_5}\left(x_4, x_5\right)}$$

$$\text{for } f_{X_4X_5}\left(x_4, x_5\right) > 0.$$

6-5: More Than Two Random Variables

6-5.5 Independence with Multiple Variables

The concept of independence can be extended to multiple variables.

Random variables $X_1, X_2, \ldots, X_p$ are independent if and only if

$$f_{X_1X_2\ldots X_p}\left(x_1, x_2, \ldots, x_p\right) = f_{X_1}\left(x_1\right) \cdot f_{X_2}\left(x_2\right) \ldots f_{X_p}\left(x_p\right) \text{ for } \mathit{all} \ \mathrm{x}_1, \mathrm{x}_2, \ldots, \mathrm{x}_p \quad (6\text{-}16)$$

(joint pdf equals the product of all the marginal PDFs)

6-5: More Than Two Random Variables

6-5.5 Independence with Multiple Variables

Example 6-16: Layer Thickness

Suppose X_1, X_2, X_3 represent the thickness in μm of a substrate, an active layer and a coating layer of a chemical product. Assume that these variables are independent and normally distributed with parameters and specified limits as tabled.

What proportion of the product meets all specifications?
Answer: 0.7783, 3 layer product.

Which one of the three thicknesses has the least probability of meeting specs?
Answer: Layer 3 has least prob.

	Normal Random Variables		
	X_1	X_2	X_3
Mean (μ)	10,000	1,000	80
Std dev (σ)	250	20	4
Lower limit	9,200	950	75
Upper limit	10,800	1,050	85
P(in limits)	0.99863	0.98758	0.78870
		P(all in limits) =	0.77783

6-6: Covariance & Correlation

6-6.3 Independence Implies ρ = 0

- If X and Y are independent random variables,

$$\sigma_{XY} = \rho_{XY} = 0$$

$\rho_{XY} = 0$ is necessary, but not a sufficient condition for independence.

  - Figure 6-10d (x, y plots as a circle) provides an example.
  - Figure 6-10b (x, y plots as a square) indicates independence, but a non-rectangular pattern would indicate dependence.

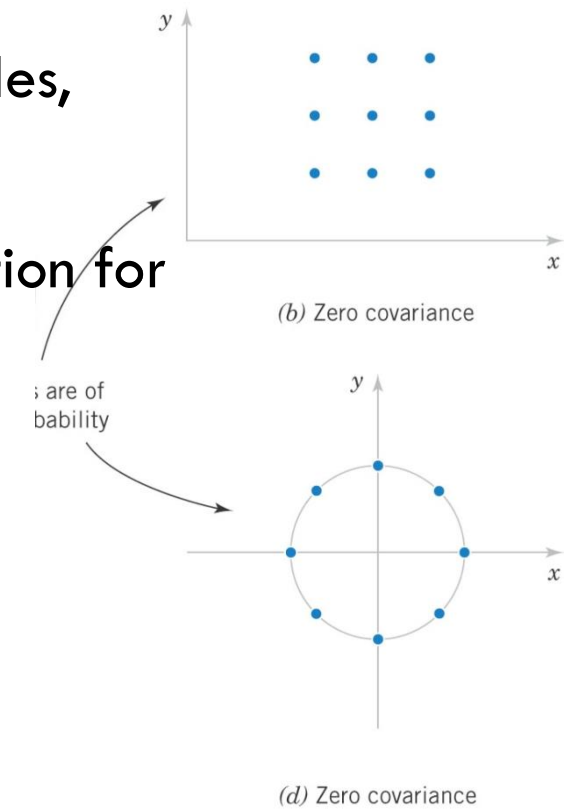


Figure 6-10 Joint probability distributions and the sign of cov(X, Y).

6-6: Covariance & Correlation

Example 6-21: Independence Implies Zero Covariance

Let $f_{XY}(xy) = x \cdot y/16$ for $0 \le x \ge 2$ and $0 \le y \le 4$

Show that $\sigma_{XY} = E(XY) - E(X) \cdot E(Y) = 0$

$$E(X) = \frac{1}{16}\int_0^4 \left[\int_0^2 x^2 y\,dx\right] dy$$

$$= \frac{1}{16}\int_0^4 y \left[\left.\frac{x^3}{3}\right|_0^2\right] dy$$

$$= \frac{1}{16}\left[\left.\frac{y^2}{2}\right|_0^4\right]\left[\frac{8}{3}\right] = \frac{1}{6}\cdot\frac{16}{2} = \frac{4}{3}$$

$$E(Y) = \frac{1}{16}\int_0^4 \left[\int_0^2 xy^2\,dx\right] dy$$

$$= \frac{1}{16}\int_0^4 y^2 \left[\left.\frac{x^2}{2}\right|_0^2\right] dy$$

$$= \frac{2}{16}\left[\left.\frac{y^3}{3}\right|_0^4\right] = \frac{1}{8}\cdot\frac{64}{3} = \frac{8}{3}$$

$$E(XY) = \frac{1}{16}\int_0^4 \left[\int_0^2 x^2 y^2\,dx\right] dy$$

$$= \frac{1}{16}\int_0^4 y^2 \left[\left.\frac{x^3}{3}\right|_0^2\right] dy$$

$$= \frac{1}{16}\int_0^4 y^2 \left[\frac{8}{3}\right] dy$$

$$= \frac{1}{6}\left[\left.\frac{y^3}{3}\right|_0^4\right] = \frac{1}{6}\cdot\frac{64}{3} = \frac{32}{9}$$

$$\sigma_{XY} = E(XY) - E(X) - E(Y)$$

$$= \frac{32}{9} - \frac{4}{3}\cdot\frac{8}{3} = 0$$

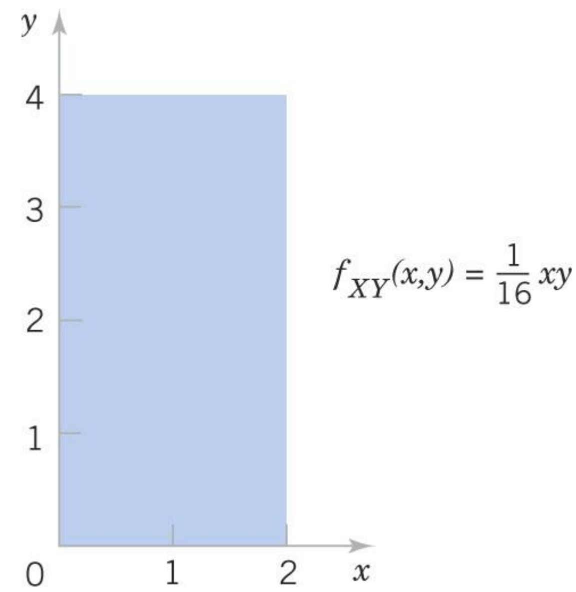


Figure 6-13
A planar joint distribution.

6-7: Common Joint Distributions

- There are two common joint distributions
  - Multinomial probability distribution (discrete), an extension of the binomial distribution
  - Bivariate normal probability distribution (continuous), a two-variable extension of the normal distribution. Although they exist, we do not deal with more than two random variables.
- There are many lesser known and custom joint probability distributions as you have already seen.

6-7: Common Joint Distributions

6-7.1 Multinomial Probability Distribution

- Suppose a random experiment consists of a series of n trials. Assume that:

  1) The outcome of each trial can be classifies into one of k classes.

  2) The probability of a trial resulting in one of the k outcomes is constant, denoted as $p_1, p_2, \ldots, p_k$.

  3) The trials are independent.

- The random variables $X_1, X_2, \ldots, X_k$ denote the number of outcomes in each class and have a multinomial distribution and probability mass function:

$$P\left(X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k\right) = \frac{n!}{x_1!x_2!\ldots x_k!} p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k} \qquad \text{(6-20)}$$

for $x_1 + x_2 + \ldots + x_k = n$ and $p_1 + p_2 + \ldots + p_k = 1$.

6-7: Common Joint Distributions

6-7.1 Multinomial Probability Distribution

Example 6-22: Digital Channel (Cont.)

Of the 20 bits received over a digital channel, 14 are of excellent quality, 3 are good, 2 are fair, 1 is poor. The sequence received was EEEEEEEEEEEEEEGGGFFP.

The probability of that sequence is $0.6^{14}0.3^{3}0.08^{2}0.02^{1} = 2.708*10^{-9}$

However, the number of different ways of receiving those bits is a lot!

x	P(x)
E	0.60
G	0.30
F	0.08
P	0.02

The combined result is a multinomial distribution.

$$P\left(x_1 = 14, x_2 = 3, x_3 = 2, x_4 = 1\right) = \frac{20!}{14!3!2!1!} 0.6^{14}0.3^{3}0.08^{2}0.02^{1} = 0.0063$$

6-7: Common Joint Distributions

6-7.1 Multinomial Probability Distribution

Example 6-22: Digital Channel (Cont.)

What is the probability that 12 bits are *E*, 6 bits are *G*, 2 are *F*, and 0 ~~are~~ is *P*?

x	P(x)
E	0.60
G	0.30
F	0.08
P	0.02

12+6+2+0 = 20

$$P\left(x_1 = 12, x_2 = 6, x_3 = 2, x_4 = 0\right) = \frac{20!}{12!6!2!0!} 0.6^{12} 0.3^{6} 0.08^{2} 0.02^{0}$$

$$= 0.0358$$

Using Excel	
0.03582	= (FACT(20)/(FACT(12)*FACT(6)*FACT(2))) * 0.6^12*0.3^6*0.08^2

6-7: Common Joint Distributions

6-7.1 Multinomial Probability Distribution

Example 6-22: Digital Channel (Cont.)

What is the probability that 12 bits are *E*, 6 bits are *G*, 2 are *F*, and 0 are *P*?

$$P\left(x_1 = 12, x_2 = 6, x_3 = 2, x_4 = 0\right) = \frac{20!}{12!6!2!0!} 0.6^{12} 0.3^{6} 0.08^{2} 0.02^{0}$$
$$= 0.0358$$

x	P(x)
E	0.60
G	0.30
F	0.08
P	0.02

Using Excel	
0.03582	= (FACT(20)/(FACT(12)*FACT(6)*FACT(2))) * 0.6^12*0.3^6*0.08^2

6-7: Common Joint Distributions

6-7.1 Multinomial Probability Distribution

Multinomial Means and Variances

The marginal distributions of the multinomial are binomial.

If $X_1, X_2, \ldots, X_k$ have a multinomial distribution, the marginal probability distributions of X_i is binomial with:

$$E(X_i) = np_i \quad \text{and} \quad V(X_i) = np_i(1-p_i) \qquad (6\text{-}21)$$

6-7: Common Joint Distributions

6-7.1 Multinomial Probability Distribution

Example 6-23: Digital Channel (Cont.)

Refer again to the prior Example 6-22.

The classes are now {G}, {F}, and {E, P}.

Now the multinomial changes to:

$$P_{X_2 X_3}(x_2, x_3) = \frac{n!}{x_2!\, x_3!\, (n - x_2 - x_3)!} p_2^{x_2} p_3^{x_3} (1 - p_2 - p_3)^{n - x_2 - x_3}$$

for $x_2 = 0$ to $n - x_3$ and $x_3 = 0$ to $n - x_2$.

6-7: Common Joint Distributions

6-7.2 Bivariate Normal Distribution

Earlier, we discussed the two dimensions of an injection-molded part as two random variables (X and Y). Let each dimension be modeled as a normal random variable. Since the dimensions are from the same part, they are typically not independent and hence correlated.

Now we have five parameters to describe the bivariate normal distribution:

$$\mu_X, \sigma_X, \mu_Y, \sigma_Y, \rho_{XY}$$

6-7: Common Joint Distributions

6-7.2 Bivariate Normal Distribution

$$f_{XY}\left(x, y; \mu_X, \sigma_X, \mu_Y, \sigma_Y, \rho\right) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} e^u$$

$$u = \frac{1}{2\left(1-\rho^2\right)}\left[\frac{\left(x-\mu_X\right)^2}{\sigma_X^2} - \frac{2\rho\left(x-\mu_X\right)\left(y-\mu_Y\right)}{\sigma_X\sigma_Y} + \frac{\left(y-\mu_Y\right)^2}{\sigma_Y^2}\right] \qquad (6\text{-}21)$$

for $-\infty < x < \infty$ and $-\infty < y < \infty$.

Parameter limits: $\begin{cases} \sigma_x > 0, & -\infty < \mu_x < \infty, \\ \sigma_y > 0, & -\infty < \mu_y < \infty, \end{cases} \quad -1 < \rho < 1$

6-7: Common Joint Distributions

6-7.2 Bivariate Normal Distribution

Role of Correlation

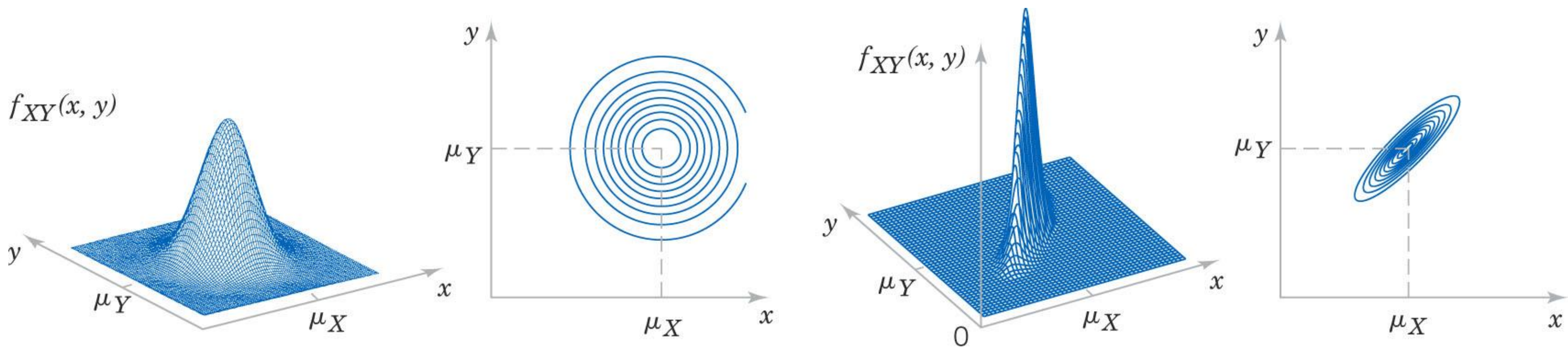


Figure 6-14 These illustrations show the shapes and contour lines of two bivariate normal distributions. The left distribution has **independent** *X*, *Y* random variables ($\rho = 0$). The right distribution has **dependent** *X*, *Y* random variables with positive correlation ($\rho > 0$, actually 0.9). The center of the contour ellipses is the point (μ_X, μ_Y).

6-7: Common Joint Distributions

6-7.2 Bivariate Normal Distribution

Standard Bivariate Normal Distribution

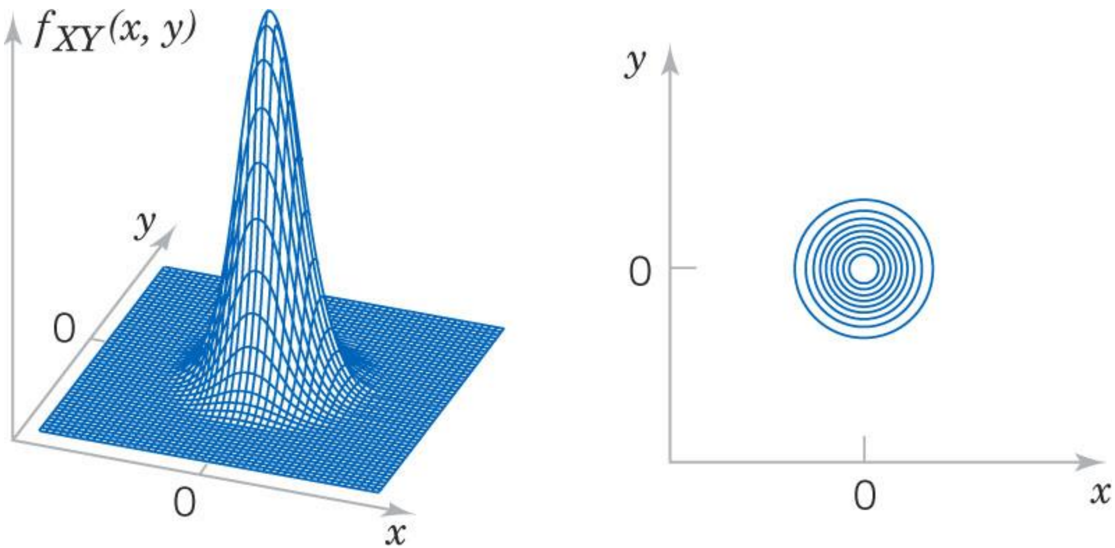


Figure 6-15 This is a standard bivariate normal because its means are zero, its standard deviations are one, and its correlation is zero since X and Y are independent. The density function is:

$$f_{XY}(x, y) = \frac{1}{\sqrt{2\pi}} e^{-0.5\left(x^2+y^2\right)} \qquad (6\text{-}22)$$

6-7: Common Joint Distributions

6-7.2 Bivariate Normal Distribution

Marginal Distributions of the Bivariate Normal

If X and Y have a bivariate normal distribution with joint probability density function $f_{XY}(x,y;\sigma_X,\sigma_Y,\mu_X,\mu_Y,\rho)$, the marginal probability distributions of X and Y are normal with means μ_X and μ_Y and σ_X and σ_Y, respectively.

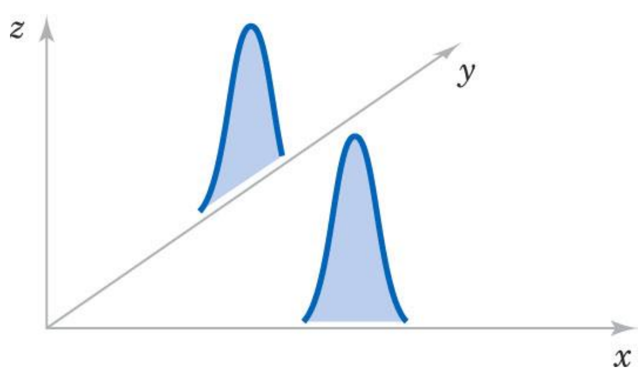


Figure 6-16 The marginal probability density functions of a bivariate normal distribution are simply projections of the joint onto each of the axis planes. Note that the correlation (ρ) has no effect on the marginal distributions.

6-7: Common Joint Distributions

6-7.2 Bivariate Normal Distribution

Conditional Distributions of the Joint Normal

If X and Y have a bivariate normal distribution with joint probability density $f_{XY}(x,y;\sigma_X,\sigma_Y,\mu_X,\mu_Y,\rho)$, the conditional probability distribution of Y given $X = x$ is normal with mean and variance as follows:

$$\mu_{Y|x} = \mu_Y - \rho\frac{\sigma_Y}{\sigma_X}\left(x - \mu_X\right) \qquad (6\text{-}23)$$

$$\sigma^2_{Y|x} = \sigma^2_Y\left(1 - \rho^2\right) \qquad (6\text{-}24)$$

6-7.2 Bivariate Normal Distribution

Correlation of Bivariate Normal Random Variables

If X and Y have a bivariate normal distribution with joint probability density function $f_{XY}(x,y;\sigma_X,\sigma_Y,\mu_X,\mu_Y,\rho)$, the correlation between X and Y is ρ.

- In general, zero correlation does not imply independence.
- But in the special case that X and Y have a bivariate normal distribution, if $\rho = 0$, then X and Y are independent.

6-7: Common Joint Distributions

6-7.2 Bivariate Normal Distribution

Example 6-24: Injection-Molded Part

The injection- molded part dimensions has parameters as tabled and is graphed as shown.

The probability of *X* and *Y* being within limits is the volume within the PDF between the limit values.

This volume is determined by numerical integration – beyond the scope of this text.

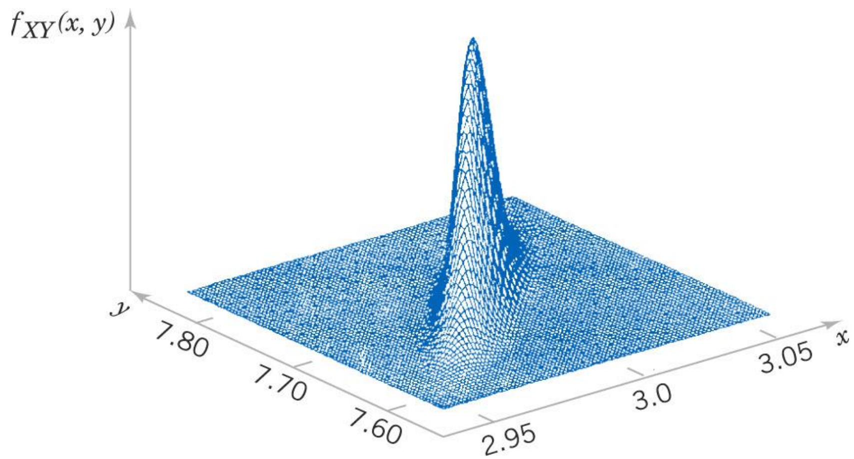


Figure 6-17

	Bivariate	
	X	Y
Mean	3	7.7
Std Dev	0.04	0.08
Correlation	0.8	
Upper Limit	3.05	7.80
Lower Limit	2.95	7.60

6-8: Linear Functions of Random Variables

- A function of random variables is itself a random variable.
- A function of random variables can be formed by either linear or nonlinear relationships. We limit our discussion here to linear functions.
- Given random variables $X_1, X_2,\ldots,X_p$ and constants $c_1, c_2, \ldots, c_p$

$$Y = c_1X_1 + c_2X_2 + \ldots + c_pX_p \qquad (6\text{-}25)$$

is a linear combination of $X_1, X_2,\ldots,X_p$.

6-8: Linear Functions of Random Variables

6-8.1 Mean & Variance of a Linear Function

Let $Y = c_1X_1 + c_2X_2 + \ldots + c_pX_p$

$$E(Y) = c_1E(X_1) + c_2E(X_2) + \ldots + c_pE(X_p)$$

$$\mathrm{V}(Y) = c_1^2V(X_1) + c_2^2V(X_2) + \ldots + c_p^2V(X_p) + 2\sum_{i<j}\sum c_ic_j \operatorname{cov}(X_iX_j)$$

If $X_1, X_2, \ldots, X_p$ are independent, then $\operatorname{cov}(X_iX_j) = 0,$

$$V(Y) = c_1^2V(X_1) + c_2^2V(X_2) + \ldots + c_p^2V(X_p)$$

6-8: Linear Functions of Random Variables

6-8.1 Mean & Variance of a Linear Function

Example 6-25: Negative Binomial Distribution

Let X_i be a geometric random variable with parameter p with

$\mu = 1/p$ and $\sigma^2 = (1-p)/p^2$

Let $Y = X_1 + X_2 + \ldots + X_r$, a linear combination of r independent geometric random variables.

Then Y is a negative binomial random variable with

$\mu = r/p$ and $\sigma^2 = r(1-p)/p^2$.

Thus, a negative binomial random variable is a sum of r identically distributed and independent geometric random variables.

6-8: Linear Functions of Random Variables

6-8.1 Mean & Variance of a Linear Function

Example 6-26: Error Propagation

A semiconductor product consists of three independent layers. The variances of the thickness of each layer is 25, 40 and 30 nm. What is the variance of the finished product?

Answer:

$$X = X_1 + X_2 + X_3$$

$$V(X) = \sum_{i=1}^{3} V(X_i) = 25 + 40 + 30 = 95 nm^2$$

$$SD(X) = \sqrt{95} = 9.747 nm$$

6-8: Linear Functions of Random Variables

6-8.2 Reproductive Property of the Normal Distribution

If $X_1, X_2, ..., X_p$ are independent, normal random variables

with $E(X_i) = \mu$, and $V(X_i) = \sigma^2$, for $i = 1, 2, ..., p$,

then $Y = c_1 X_1 + c_2 X_2 + ... + c_p X_p$

is a normal random variable with

$$E(Y) = c_1 \mu_1 + c_2 \mu_2 + ... + c_p \mu_p$$

and

$$V(Y) = c_1^2 \sigma_1^2 + c_2^2 \sigma_2^2 + ... + c_p^2 \sigma_p^2$$

6-8: Linear Functions of Random Variables

6-8.2 Reproductive Property of the Normal Distribution

Example 6-27: Linear Function of Independent Normals

Let the random variables X_1 and X_2 denote the independent length and width of a rectangular manufactured part. Their parameters are shown in the table.

What is the probability that the perimeter exceeds 14.5 cm?

	Parameters of	
	X_1	X_2
Mean	2	5
Std Dev	0.1	0.2

Let $Y = 2X_1 + 2X_2 =$ perimeter

$$E(Y) = 2E(X_1) + 2E(X_2) = 2(2) + 2(5) = 14 \text{ cm}$$

$$V(Y) = 2^2 V(X_1) + 2^2 V(X_2) = 4(0.1)^2 + 4(0.2)^2 = 0.04 + 0.16 = 0.20$$

$$SD(Y) = \sqrt{0.20} = 0.4472 \text{ cm}$$

$$P(Y > 14.5) = 1 - \Phi\left(\frac{14.5 - 14}{.4472}\right) = 1 - \Phi(1.1180) = 0.1318$$

Using Excel	
0.1318	= 1 - NORMDIST(14.5, 14, SQRT(0.2), TRUE)

6-8: Linear Functions of Random Variables

6-8.2 Reproductive Property of the Normal Distribution

Example 6-28: Beverage Volume

Soft drink cans are filled by an automated filling machine. The mean fill volume is 12.1 fluid ounces, and the standard deviation is 0.1 fl oz. Assume that the fill volumes are independent, normal random variables. What is the probability that the average volume of 10 cans is less than 12 fl oz?

Let X_i denote the fill volume of the i$^{\text{th}}$ can

$$\text{Let } \bar{X} = \sum_{i=1}^{10} X_i \Big/ 10$$

$$E(\bar{X}) = \sum_{i=1}^{n} E(X_i) \Big/ 10 = \frac{10 \cdot 12.1}{10} = 12.1 \text{ fl oz}$$

$$V(\bar{X}) = \left(1/10\right)^2 \sum_{i=1}^{10} V(X_i) = \frac{10(0.1)^2}{100} = 0.001 \text{ fl oz}^2$$

$$P(\bar{X} < 12) = \Phi\left(\frac{12 - 12.1}{\sqrt{0.001}}\right) = \Phi(-3.16) = 0.00079$$

Using Excel
0.000783
= NORMDIST(12, 12.1, SQRT(0.001), TRUE)

6-9: Random Samples, Statistics, and The Central Limit Theorem

6-9.1 Central Limit Theorem

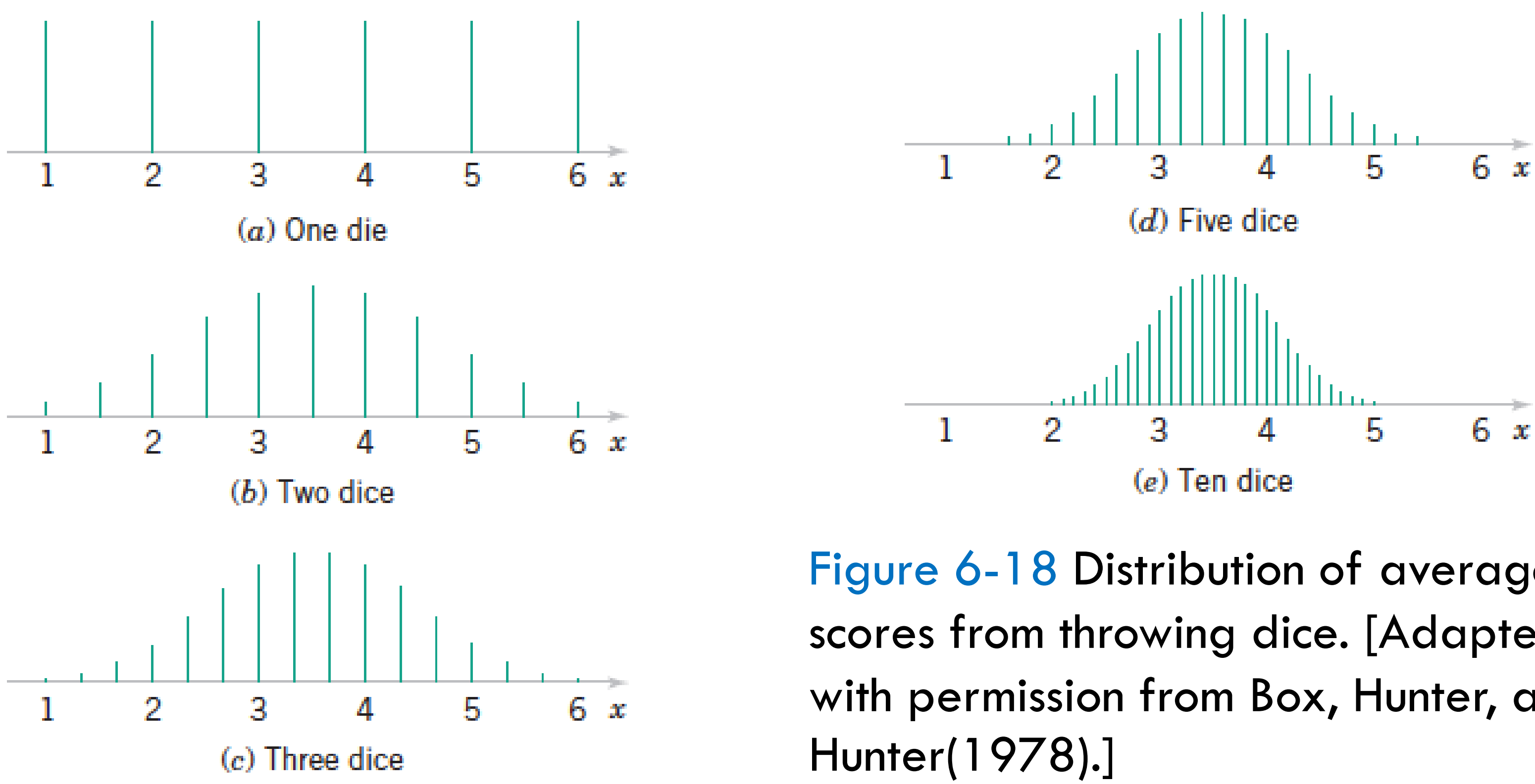


Figure 6-18 Distribution of average scores from throwing dice. [Adapted with permission from Box, Hunter, and Hunter(1978).]

6-9: Random Samples, Statistics, and The Central Limit Theorem

6-9.1 Central Limit Theorem

If $X_1, X_2, \ldots, X_n$ is a random sample of size n taken from a population with mean μ and variance σ^2, and if $\overline{X}$ is the sample mean, the limiting form of the distribution of

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \tag{6-26}$$

as $n \to \infty$, is the standard normal distribution.

6-9.2 Central Limit Theorem

Mean & Variance of an Average

If $\bar{X} = \dfrac{\left(X_1 + X_2 + \ldots + X_p\right)}{p}$ and $E\left(X_i\right) = \mu$

Then $E\left(\bar{X}\right) = \dfrac{p \cdot \mu}{p} = \mu$ (6-27)

If the X_i are independent with $V\left(X_i\right) = \sigma^2$

Then $V\left(\bar{X}\right) = \dfrac{p \cdot \sigma^2}{p^2} = \dfrac{\sigma^2}{p}$ (6-28)

6-9: Random Samples, Statistics, and The Central Limit Theorem

6-9.1 Central Limit Theorem

Example 6-29:

An electronics company manufactures resistors that have a mean resistance of $100\ \Omega$ and a standard deviation of 10Ω. Find the probability that a random sample of $n = 25$ resistors will have an average resistance less than 95Ω.

Note that the sampling distribution of $\bar{X}$ is approximately normal, with mean $\mu_{\bar{X}} = 100\Omega$ and a standard deviation of $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{25}} = 2$

Therefore, the desired probability corresponds to the shaded area in Fig. 6-19. Standardizing the point $\bar{X} = 95$, we find that $z = \frac{95-100}{2} = -2.5$

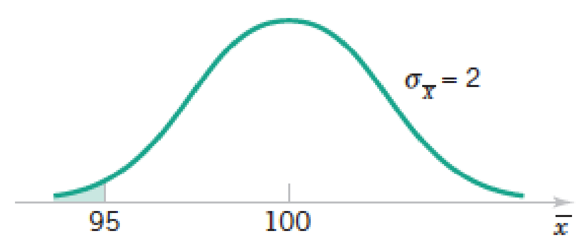


Figure 6-19 Probability density function of average resistance.

And, therefore, $P(\bar{X} < 95) = P(Z < -2.5) = 0.0062$