Chapter 5: JOINT PROBABILITY DISTRIBUTIONS

Part 1: Sections 5.1 & 5.2

For both *discrete* and *continuous* random variables we will discuss the following...

- Joint Distributions (for two or more $r.v.$'s)

- Marginal Distributions
  (computed from a joint distribution)

- Conditional Distributions
  (e.g. $P(Y = y|X = x)$)

- Independence for $r.v.$'s X and Y

This is a good time to refresh your memory on double-integration. We will be using this skill in the upcoming lectures.

Recall a discrete probability distribution (or *pmf*) for a single *r.v.* X with the example below...

x	0	1	2
$f(x)$	0.50	0.20	0.30

Sometimes we're simultaneously interested in two or more variables in a random experiment. We're looking for a relationship between the two variables.

Examples for discrete *r.v.*'s

- Year in college vs. Number of credits taken
- Number of cigarettes smoked per day vs. Day of the week

Examples for continuous *r.v.*'s

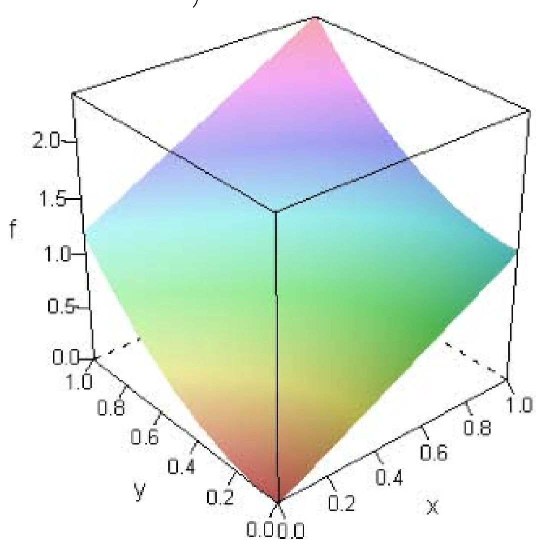
- Time when bus driver picks you up vs. Quantity of caffeine in bus driver's system
- Dosage of a drug (ml) vs. Blood compound measure (percentage)

In general, if X and Y are two random variables, the probability distribution that defines their simultaneous behavior is called a joint probability distribution.

Shown here as a table for two discrete random variables, which gives $P(X = x, Y = y)$.

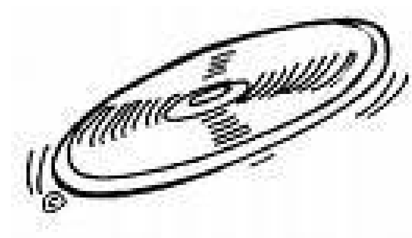
		x		
		1	2	3
	1	0	1/6	1/6
y	2	1/6	0	1/6
	3	1/6	1/6	0

Shown here as a graphic for two continuous random variables as $f_{X,Y}(x, y)$.



If X and Y are discrete, this distribution can be described with a joint probability mass function.

If X and Y are continuous, this distribution can be described with a joint probability density function.



- **Example**: Plastic covers for CDs (Discrete joint pmf)

  Measurements for the length and width of a rectangular plastic covers for CDs are rounded to the nearest mm (so they are discrete).

  Let X denote the length and
  Y denote the width.

  The possible values of X are 129, 130, and 131 mm. The possible values of Y are 15 and 16 mm (Thus, both X and Y are discrete).

There are 6 possible pairs (X, Y).

We show the probability for each pair in the following table:

		x=length		
		129	130	131
y=width	15	0.12	0.42	0.06
	16	0.08	0.28	0.04

The sum of all the probabilities is 1.0.

The combination with the highest probability is $(130, 15)$.

The combination with the lowest probability is $(131, 16)$.

The joint probability mass function is the function $f_{XY}(x, y) = P(X = x, Y = y)$. For example, we have $f_{XY}(129, 15) = 0.12$.

If we are given a joint probability distribution for X and Y, we can obtain the individual probability distribution for X or for Y (and these are called the **Marginal Probability Distributions**)...

- **Example**: Continuing plastic covers for CDs

  Find the probability that a CD cover has length of $129mm$ (i.e. $X = 129$).

x= length

		129	130	131
y=width	15	**0.12**	0.42	0.06
	16	**0.08**	0.28	0.04

$$\begin{aligned} P(X = 129) = &\ P(X = 129 \text{ and } Y = 15) \\ &+ P(X = 129 \text{ and } Y = 16) \\ = &\ 0.12 \ + \ 0.08 = 0.20 \end{aligned}$$

What is the probability distribution of X?

		x= length		
		129	130	131
y=width	15	0.12	0.42	0.06
	16	0.08	0.28	0.04
column totals		**0.20**	**0.70**	**0.10**

The probability distribution for X appears in the column totals...

x	129	130	131
$f_X(x)$	0.20	0.70	0.10

* NOTE: We've used a subscript X in the probability mass function of X, or $f_X(x)$, for clarification since we're considering more than one variable at a time now.

We can do the same for the Y random variable:

		x= length			**row totals**
		129	130	131	
y=width	15	0.12	0.42	0.06	**0.60**
	16	0.08	0.28	0.04	**0.40**
column totals		**0.20**	**0.70**	**0.10**	**1**

y	15	16
$f_Y(y)$	0.60	0.40

Because the the probability mass functions for X and Y appear in the margins of the table (i.e. column and row totals), they are often referred to as the **Marginal Distributions** for X and Y.

When there are two random variables of interest, we also use the term **bivariate probability distribution** or **bivariate distribution** to refer to the joint distribution.

- **Joint Probability Mass Function**
  The joint probability mass function of the discrete random variables X and Y, denoted as $f_{XY}(x,y)$, satisfies

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

$$(2)\ \ \sum_x \sum_y f_{XY}(x,y) = 1$$

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

For when the r.v.'s are discrete.

(Often shown with a 2-way table.)

		x= length		
		129	**130**	**131**
y=width	**15**	0.12	0.42	0.06
	16	0.08	0.28	0.04

- **Marginal Probability Mass Function**
  If X and Y are discrete random variables with joint probability mass function $f_{XY}(x, y)$, then the marginal probability mass functions of X and Y are

$$f_X(x) = \sum_y f_{XY}(x, y)$$

and

$$f_Y(y) = \sum_x f_{XY}(x, y)$$

  where the sum for $f_X(x)$ is over all points in the range of (X, Y) for which $X = x$ and the sum for $f_Y(y)$ is over all points in the range of (X, Y) for which $Y = y$.

We found the marginal distribution for X in the CD example as...

x	129	130	131
$f_X(x)$	0.20	0.70	0.10

HINT: When asked for $E(X)$ or $V(X)$ (i.e. values related to only 1 of the 2 variables) but you are given a joint probability distribution, first calculate the marginal distribution $f_X(x)$ and work it as we did before for the univariate case (i.e. for a single random variable).

- **Example**: Batteries
  Suppose that 2 batteries are randomly chosen without replacement from the following group of 12 batteries:

  3 new
  4 used (working)
  5 defective

  Let X denote the number of new batteries chosen.

  Let Y denote the number of used batteries chosen.

a) Find $f_{XY}(x,y)$
$\{$i.e. the joint probability distribution$\}$.

b) Find $E(X)$.

ANS:

a) Though X can take on values 0, 1, and 2, and Y can take on values 0, 1, and 2, when we consider them jointly, $X+Y \leq 2$. So, not all combinations of (X,Y) are possible.

There are 6 possible cases...

CASE: no new, no used (so all defective)

$$f_{XY}(0,0) = \frac{\binom{5}{2}}{\binom{12}{2}} = 10/66$$

CASE: no new, 1 used

$$f_{XY}(0,1) = \frac{\binom{4}{1}\binom{5}{1}}{\binom{12}{2}} = 20/66$$

CASE: no new, 2 used

$$f_{XY}(0,2) = \frac{\binom{4}{2}}{\binom{12}{2}} = 6/66$$

CASE: 1 new, no used

$$f_{XY}(1,0) = \frac{\binom{3}{1}\binom{5}{1}}{\binom{12}{2}} = 15/66$$

CASE: 2 new, no used

$$f_{XY}(2,0) = \frac{\binom{3}{2}}{\binom{12}{2}} = 3/66$$

CASE: 1 new, 1 used

$$f_{XY}(1,1) = \frac{\binom{3}{1}\binom{4}{1}}{\binom{12}{2}} = 12/66$$

The joint distribution for X and Y is...

x= number of *new* chosen

		0	1	2
y=number of	0	10/66	15/66	3/66
used	1	20/66	12/66	
chosen	2	6/66		

There are 6 possible (X, Y) pairs.
And, $\sum_x \sum_y f_{XY}(x,y) = 1$.

b) Find $E(X)$.

- **Joint Probability Density Function**

  A joint probability density function for the continuous random variable X and Y, denoted as $f_{XY}(x,y)$, satisfies the following properties:

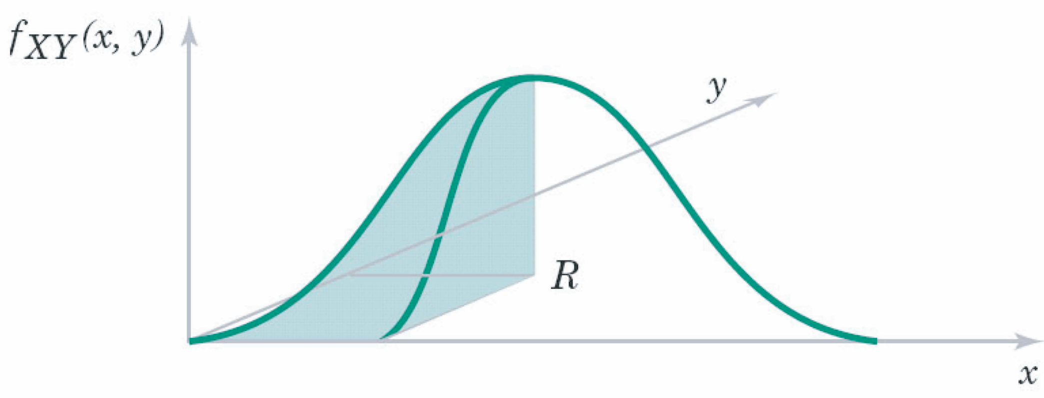
  1. $f_{XY}(x,y) \geq 0$ for all x, y

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

  3. For any region R of 2-D space

$$P((X,Y) \in R) = \int \int_R f_{XY}(x,y)\, dx\, dy$$

For when the r.v.'s are continuous.



- **Example**: Movement of a particle

  An article describes a model for the movement of a particle. Assume that a particle moves within the region A bounded by the x axis, the line $x = 1$, and the line $y = x$. Let (X, Y) denote the position of the particle at a given time. The joint density of X and Y is given by

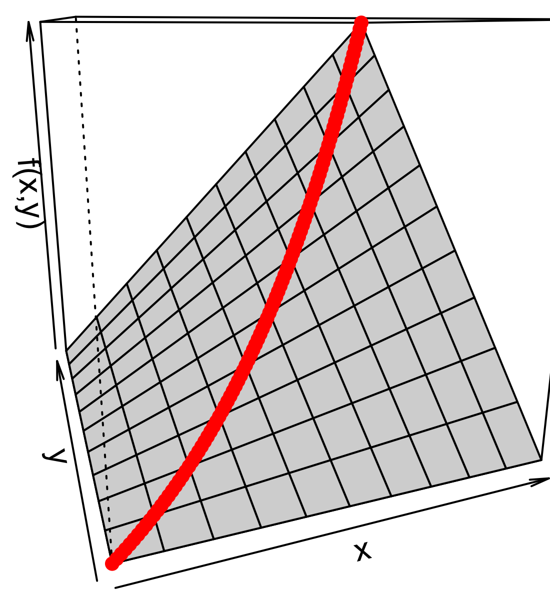
$$f_{XY}(x, y) = 8xy \quad \text{for} \quad (x, y) \in A$$

  a) Graphically show the region in the XY plane where $f_{XY}(x, y)$ is nonzero.

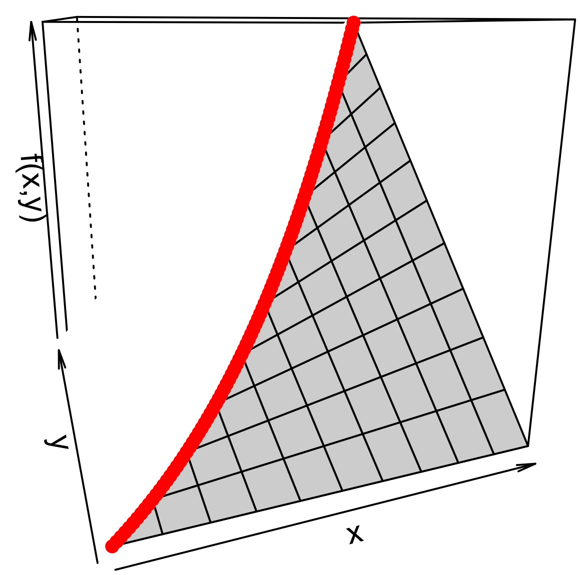
The probability density function $f_{XY}(x,y)$ is shown graphically below.

Without the information that $f_{XY}(x,y) = 0$ for (x,y) outside of A, we could plot the full surface, but the particle is only found in the given triangle A, so the joint probability density function is shown on the right.

This gives a volume under the surface that is above the region A equal to 1.



Not a *pdf*



A *pdf*

b) Find $P(0.5 < X < 1, 0 < Y < 0.5)$

c) Find $P(0 < X < 0.5, 0 < Y < 0.5)$

d) Find $P(0.5 < X < 1, 0.5 < Y < 1)$

- **Marginal Probability Density Function**

  If X and Y are continuous random variables with joint probability density function $f_{XY}(x, y)$, then the marginal density functions for X and Y are

$$f_X(x) = \int_y f_{XY}(x, y)\, dy$$

  and

$$f_Y(y) = \int_x f_{XY}(x, y)\, dx$$

  where the first integral is over all points in the range of (X, Y) for which $X = x$, and the second integral is over all points in the range of (X, Y) for which $Y = y$.

- **Mean from a Joint Distribution**

If X and Y are continuous random variables with joint probability density function $f_{XY}(x,y)$, then

$$E(X) = \int_{-\infty}^{\infty} x f_X(x)\,dx$$

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

HINT: $E(X)$ and $V(X)$ can be obtained by first calculating the marginal probability distribution of X, or $f_X(x)$.

- **Example**: Movement of a particle

An article describes a model for the movement of a particle. Assume that a particle moves within the region A bounded by the x axis, the line $x = 1$, and the line $y = x$. Let (X, Y) denote the position of the particle at a given time. The joint density of X and Y is given by

$$f_{XY}(x, y) = 8xy \quad \text{for} \quad (x, y) \in A$$

a) Find $E(X)$

Conditional Probability Distributions

Recall for events A and B,

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

We now apply this conditioning to random variables X and Y...

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

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

The conditional probability can be stated as the *joint* probability over the *marginal* probability.

Note: we can define $f_{X|y}(x)$ in a similar manner if we are interested in that conditional distribution.

- **Example**: Continuing the plastic covers...

		x= length			row totals
		129	130	131	
y=width	15	0.12	0.42	0.06	0.60
	16	0.08	0.28	0.04	0.40
column totals		0.20	0.70	0.10	1

a) Find the probability that a CD cover has a length of 130mm GIVEN the width is 15mm.

ANS: $P(X = 130|Y = 15) = \frac{P(X=130,Y=15)}{P(Y=15)}$

$$= \frac{0.42}{0.60} = 0.70$$

b) Find the conditional distribution of X given Y=15.

$$P(X = 129|Y = 15) = 0.12/0.60 = 0.20$$
$$P(X = 130|Y = 15) = 0.42/0.60 = 0.70$$
$$P(X = 131|Y = 15) = 0.06/0.60 = 0.10$$

Once you're GIVEN that Y=15, you're in a 'different space'.

		x= length			row totals
		129	130	131	
y=width	15	0.12	0.42	0.06	0.60
	16	0.08	0.28	0.04	0.40
column totals		0.20	0.70	0.10	1

For the subset of the covers with a width of $15mm$, how are the lengths (X) distributed.

The conditional distribution of X given Y=15, or $f_{X|Y=15}(x)$:

x	129	130	131
$f_{X\|Y=15}(x)$	0.20	0.70	0.10

The sum of these probabilities is 1, and this is a legitimate probability distribution .

* NOTE: Again, we use the subscript $X|Y$ for clarity to denote that this is a conditional distribution.

A conditional probability distribution $f_{Y|x}(y)$ has the following properties are satisfied:

- **For discrete random variables (X,Y)**

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

  $$(2)\quad \sum_{y} f_{Y|x}(y) = 1$$

  $$(3)\quad f_{Y|x}(y) = P(Y = y|X = x)$$

- **For continuous random variables (X,Y)**

  1. $f_{Y|x}(y) \geq 0$
  2. $\int_{-\infty}^{\infty} f_{Y|x}(y)\, dy = 1$
  3. $P(Y \in B|X = x) = \int_B f_{Y|x}(y)\, dy$ for any set B in the range of Y

- **Conditional Mean and Variance for DISCRETE random variables**

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

$$E(Y|x) = \sum_y y f_{Y|X}(y) = \mu_{Y|x}$$

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

$$V(Y|x) = \sum_y (y - \mu_{Y|x})^2 f_{Y|X}(y)$$

$$= \sum_y y^2 f_{Y|X}(y) - \mu^2_{Y|x}$$

$$= E(Y^2|x) - [E(Y|x)]^2$$

$$= \sigma^2_{Y|x}$$

• **Example**: Continuing the plastic covers...

		x=length			row totals
		129	130	131	
y=width	15	0.12	0.42	0.06	0.60
	16	0.08	0.28	0.04	0.40
column totals		0.20	0.70	0.10	1

a) Find the $E(Y|X = 129)$ and $V(Y|X = 129)$.

ANS:
We need the conditional distribution first...

y	15	16
$f_{Y\|X=129}(y)$		

- **Conditional Mean and Variance for CONTINUOUS random variables**

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

$$E(Y|x) = \int y f_{Y|x}(y)\, dy$$

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

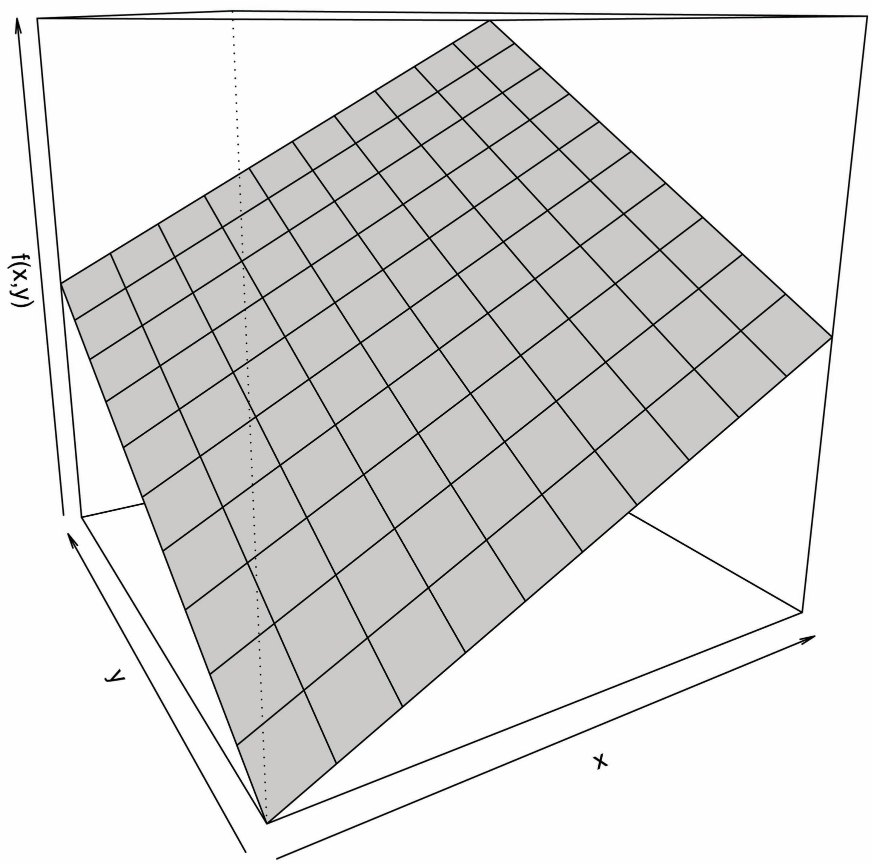
$$V(Y|x) = \int (y - \mu_{Y|x})^2 f_{Y|x}(y)\, dy$$

$$= \int y^2 f_{Y|x}(y)\, dy - \mu^2_{Y|x}$$

- **Example 1**: Conditional distribution

Suppose (X, Y) has a probability density function...

$f_{XY}(x, y) = x + y$ for $0 < x < 1, 0 < y < 1$



a) Find $f_{Y|x}(y)$.

b) Show $\int_{-\infty}^{\infty} f_{Y|x}(y) dy = 1$.

a)

b)

One more...

c) What is the conditional mean of Y given $X = 0.5$?

ANS:

First get $f_{Y|X=0.5}(y)$

$$f_{Y|x}(y) = \frac{x+y}{x+0.5} \quad \text{for } 0 < x < 1 \text{ and } 0 < y < 1$$

$$f_{Y|X=0.5}(y) = \frac{0.5+y}{0.5+0.5} = 0.5 + y \quad \text{for} \quad 0 < y < 1$$

$$E(Y|X=0.5) = \int_0^1 y(0.5+y)\,dy = \frac{7}{12}$$

Independence

As we saw earlier, sometimes, knowledge of one event does not give us any information on the probability of another event.

Previously, we stated that if A and B were independent, then

$$P(A|B) = P(A).$$

In the framework of probability distributions, if X and Y are independent random variables, then $f_{Y|X}(y) = f_Y(y)$.

- **Independence**
  For random variables X and Y, if any of the following properties is true, the others are also true, and X and Y are <u>independent</u>.

  (1) $f_{XY}(x, y) = f_X(x) f_Y(y)$ for all x and y

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

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

  (4) $P(X \in A, Y \in B) = P(X \in A) \cdot P(Y \in B)$
  for any sets A and B in the range of X and Y.

Notice how (1) leads to (2):

$$f_{Y|x}(y) = \frac{f_{XY}(x,y)}{f_X(x)} = \frac{f_X(x) f_Y(y)}{f_X(x)} = f_Y(y)$$

- **Example 1: (discrete)**
  Continuing the battery example

  Two batteries were chosen without replacement.

  Let X denote the number of new batteries chosen.

  Let Y denote the number of used batteries chosen.

		x= number of *new* chosen		
		0	1	2
y=number	0	10/66	15/66	3/66
of used	1	20/66	12/66	
chosen	2	6/66		

a) Without doing any calculations, can you tell whether X and Y are independent?

- **Example 2: (discrete)**
  Independent random variables

  Consider the random variables X and Y, which both can take on values of 0 and 1.

		x		row totals
		0	1	
y	0	0.08	0.02	0.10
	1	0.72	0.18	0.90
column totals		0.80	0.20	1

a) Are X and Y independent?

y	0	1
$f_{Y\|X=0}(y)$		

y	0	1
$f_{Y\|X=1}(y)$		

Does $f_{Y|x}(y) = f_Y(y)$ for all x & y?

Does $f_{XY}(x, y) = f_X(x) f_Y(y)$ for all x & y?

		x: 0	x: 1	row totals
y	0	0.08	0.02	0.10
	1	0.72	0.18	0.90
column totals		0.80	0.20	1

i.e. Does $P(X = x, Y = y)$
$= P(X = x) \cdot P(Y = y)$?

- **Example 3: (continuous)**
  Dimensions of machined parts (Example 5-12).

  Let X and Y denote the lengths of two dimensions of a machined part.

  X and Y are independent and measured in millimeters (you're given independence here).

  $X \sim N(10.5, 0.0025)$
  $Y \sim N(3.2, 0.0036)$

  a) Find
  $P(10.4 < X < 10.6,\ 3.15 < Y < 3.25)$.