MA 6.101
Probability and Statistics

Tejas Bodas

Assistant Professor, IIIT Hyderabad

Logistics

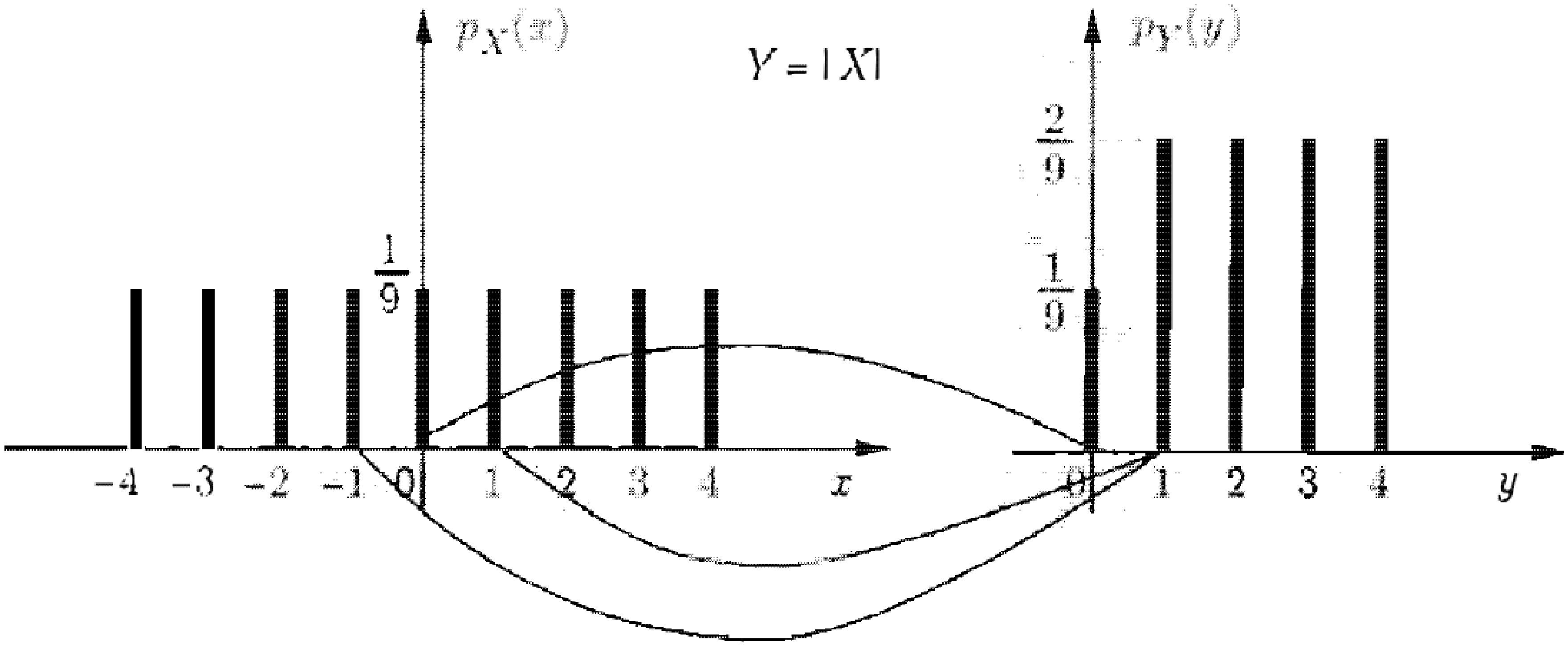
- Tutorial has been uploaded.
- Practice a lot of problems from Bertsekas.
- Quiz 1 has been corrected.
- Paper distribution tomorrow at 3:30-5pm ?
- Venue will be emailed later.

Function of random variables

- Consider $Y = |X|$ where X is the outcome of an experiment where an integer is chosen uniformly from -4 to 4.
- $p_X(x) = \frac{1}{9}$ for $x \in \{-4, -3, \ldots, 3, 4\}$.
- What is the range Ω' for Y? $\Omega' = \{0, \ldots, 4\}$.
- What is $p_Y(2)$?
- $p_Y(2) = \sum_{\{x:|x|=2\}} p_X(x) = p_X(-2) + p_x(2) = \frac{2}{9}$.

Function of random variables

- $p_Y(2) = \sum_{\{x:|x|=2\}} p_X(x) = p_X(-2) + p_x(2) = \frac{2}{9}$.



Suppose $Y = g(X)$ and X is discrete with pmf $p_X(\cdot)$. Then $p_Y(y) = \sum_{\{x:g(x)=y\}} p_X(x)$.

Function of random variables

Suppose $Y = g(X)$ and X is discrete with pmf $p_X(\cdot)$. Then $p_Y(y) = \sum_{\{x:g(x)=y\}} p_X(x)$.

Proof:

- Then what is $p_Y(y)$?
- $p_Y(y) = \mathbb{P}\{\omega \in \Omega : Y(\omega) = y\}$.
- $p_Y(y) = \mathbb{P}\{\omega \in \Omega : g(X(\omega)) = y\}$.
- $p_Y(y) = \mathbb{P}\{\omega \in \Omega : X(\omega) = g^{-1}(y)\}$.
- Is there a problem with the text in red?
- Is $g^{-1}(y)$ a value or a set?

Function of random variables

Suppose $Y = g(X)$ and X is discrete with pmf $p_X(\cdot)$. Then $p_Y(y) = \sum_{\{x:g(x)=y\}} p_X(x)$.

Proof Continued:

- $g^{-1}(y)$ is a value if $g(\cdot)$ is one to one.
- If $g(\cdot)$ is many to one, then $g^{-1}(y) := \{x : g(x) = y\}$.
- In that case, $p_Y(y) = \mathbb{P}\{\omega \in \Omega : X(\omega) \in g^{-1}(y)\}$.
- $p_Y(y) = \mathbb{P}\{\omega \in \Omega : X(\omega) \in \{x : g(x) = y\}\}$.
- Now $\mathbb{P}\{\omega \in \Omega : X(\omega) \in B\} = \sum_{\{x \in B\}} p_X(x)$ for $B \in \mathcal{F}'$.
- Proof follows after setting $B = \{x : g(x) = y\}$ $\square$

$E[g(X)]$

Suppose $Y = g(X)$ and X is discrete with pmf $p_X(\cdot)$. Then, $E[Y] = \sum_x g(x)p_X(x)$

Proof

$$
\begin{aligned}
E[Y] &= \sum_y y p_Y(y) \\
&= \sum_y \sum_{\{x: g(x)=y\}} g(x)p_X(x) \\
&= \sum_x g(x)p_X(x).
\end{aligned}
$$

$\square$

Function of continuous random variables

- Consider $Y = aX + b$ where X is a continuous random variable.
- What is $F_Y(y)$ and $f_Y(y)$?
- $F_Y(y) = P(Y \leq y) = P(aX + b \leq y)$.
- $F_Y(y) = F_X(\frac{y-b}{a})$ if $a > 0$
- $f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{1}{a} f_X(\frac{y-b}{a})$ when $a > 0$
- $F_Y(y) = 1 - F_X(\frac{y-b}{a})$ if $a < 0$
- $f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{-1}{a} f_X(\frac{y-b}{a})$ when $a < 0$
- In general, $f_Y(y) = \frac{1}{|a|} f_X(\frac{y-b}{a})$

Function of continuous random variables

Consider $Y = aX + b$ where X is a continuous random variable. Then $f_Y(y) = \frac{1}{|a|} f_X(\frac{y-b}{a})$.

- What if $Y = g(X)$ where $g(\cdot)$ is continuous, differentiable and monotone. Any guess?
- Since $g(.)$ is monotone and continuous it is invertible. Let $h(.)$ denote the inverse function. Then $h(Y) = X$.

Consider $Y = g(X)$ where g is monotone, continuous, differentiable. Then $f_Y(y) = |\frac{dh}{dy}(y)| f_X(h(y))$ where h is the inverse function of g.

Function of continuous random variables

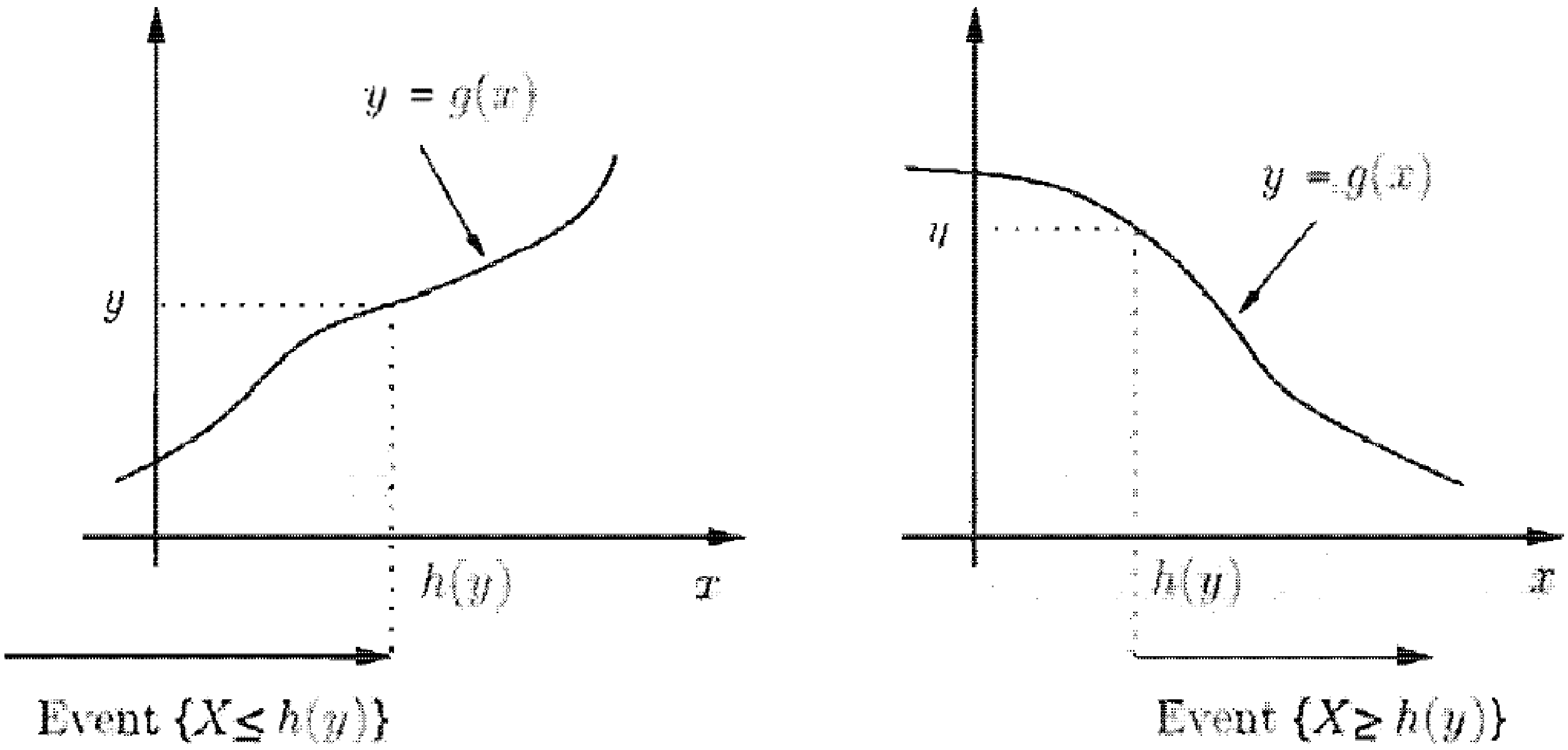
Consider $Y = g(X)$ where g is monotone, continuous, differentiable. Then $f_Y(y) = f_X(h(y))|\frac{dh}{dy}(y)|$ where h is the inverse function of g.

Proof:

- Since $g(.)$ is monotone and continuous it is invertible. Let $h(.)$ denote the inverse function. Then $X = h(Y)$.
- $F_Y(y) = P(g(X) \leq y)$.
- Is $P(g(X) \leq y) = P(X \leq h(y))$ always?
- Are the two events $\{g(X) \leq y\}$ and $X \leq h(y)$ same?
- If they are same, then the two probabilities are equal.

Function of continuous random variables

- Are the two events $\{g(X) \leq y\}$ and $\{X \leq h(y)\}$ same ?



- Same when g is increasing and compliments when g is decreasing.

Function of continuous random variables

- Are the two events $\{g(X) \leq y\}$ and $\{X \leq h(y)\}$ same ?
- Same when g is increasing and compliments when g is decreasing.
- CASE 1: $g(x)$ is non-decreasing
- $F_Y(y) = P(g(X) \leq y) = P(X \leq h(y)) = F_X(h(y))$.
- $f_Y(y) = \frac{d}{dy}(F_X(h(y))) = f_X(h(y))\frac{dh}{dy}(y)$ where $\frac{dh}{dy}(y) \geq 0$ as h is also non-decreasing.
- Rewritten therefore as $f_Y(y) = f_X(h(y))|\frac{dh}{dy}(y)|$

Function of continuous random variables

- Are the two events $\{g(X) \leq y\}$ and $\{X \leq h(y)\}$ same ?
- Same when g is increasing and compliments when g is decreasing.
- CASE 2: $g(x)$ is non-increasing
- $F_Y(y) = P(g(X) \leq y) = P(X > h(y)) = 1 - F_X(h(y))$.
- $f_Y(y) = -\frac{d}{dy}(F_X(h(y))) = -f_X(h(y))\frac{dh}{dy}(y)$ where $\frac{dh}{dy}(y) \leq 0$ as h is non-increasing as well.
- Rewritten therefore as $f_Y(y) = f_X(h(y))|\frac{dh}{dy}(y)|$. $\square$

HW: What about the case when g is not monotone ?

Multiple random variables

A running example

- Consider an experiment of tossing a coin and a dice together.
- $\Omega = \{0, 1\} \times \{1, 2, 3, 4, 5, 6\}$. $\mathcal{F} = 2^{\Omega}$. $\mathbb{P}(\omega) = \frac{1}{12}$.
- Let X and Y denote the random variables depicting outcome of a coin and dice respectively.
- For $\omega = (1, 5)$ we have $X(\omega) = 1$ and $Y(\omega) = 5$.
- We are now interested in the joint PMF $p_{XY}(x, y)$ and joint CDF $F_{XY}(x, y)$ of X and Y together.

An example

- We are now interested in the joint PMF $p_{XY}(x, y)$ and joint CDF $F_{XY}(x, y)$ of X and Y together.
- $p_{XY}(x, y) := \mathbb{P}\{\omega \in \Omega : X(w) = x \text{ and } Y(\omega) = y\}$.
- $F_{XY}(x, y) := \mathbb{P}\{\omega \in \Omega : X(w) \leq x \text{ and } Y(\omega) \leq y\}$.
- We can use PMF to calculate $P((X, Y) \in A)$.
- $P((X, Y) \in A) = \mathbb{P}\{\omega \in \Omega : (X(\omega), Y(\omega)) \in A\}$. Therefore $P((X, Y) \in A) = \sum_{(x,y)\in A} p_{XY}(x, y)$.
- Suppose A is the event that you get a head and the roll is even. What is $P((X, Y) \in A)$?

Marginals

- What is $p_{XY}(1, i)$? $(= \frac{1}{12})$.
- $\sum_i p_{XY}(1, i) = \mathbb{P}\{\omega \in \Omega : X(\omega) = 1\} = \frac{1}{2} = p_X(x)$.
- Similarly, $p_{XY}(1, i) + p_{XY}(0, i) = \frac{1}{6} = p_Y(i)$.

> The marginal PMF's p_X and p_Y can be obtained from the joint PMF as follows:
> $p_X(x) = \sum_y p_{XY}(x, y)$ and $p_Y(y) = \sum_x p_{XY}(x, y)$.

This is true in general, and requires a proof (later).

Recap

- Functions of random variables ($Y = g(X)$)
- Obtain $F_Y(y)$ and $f_Y(y)$ from $F_X(x)$ and $f_X(x)$.
- Special case when $g(.)$ is monotone.
- Joint PMF $p_{XY}(x, y)$ and CDF $F_{XY}(x, y)$.
- $p_{XY}(x, y) := \mathbb{P}\{\omega \in \Omega : X(w) = x \text{ and } Y(\omega) = y\}$.
- $F_{XY}(x, y) := \mathbb{P}\{\omega \in \Omega : X(w) \leq x \text{ and } Y(\omega) \leq y\}$.
- The marginal PMF's p_X and p_Y can be obtained from the joint PMF as follows:
  $p_X(x) = \sum_y p_{XY}(x, y)$ and $p_Y(y) = \sum_x p_{XY}(x, y)$.

Marginals

The marginal PMF's p_X and p_Y can be obtained from the joint PMF as follows:
$p_X(x) = \sum_y p_{XY}(x, y)$ and $p_Y(y) = \sum_x p_{XY}(x, y)$.

Proof:

$$
\begin{aligned}
P_X(x) &= \mathbb{P}\{\omega \in \Omega : X(\omega) = x\} \\
&= \mathbb{P}\{\bigcup_y \{\omega \in \Omega : X(\omega) = x, Y(\omega) = y\}\} \\
&= \sum_y \mathbb{P}\{\{\omega \in \Omega : X(\omega) = x, Y(\omega) = y\}\}
\end{aligned}
$$

Independence

- Back with the running example of coin and dice.
- Write down $p_{XY}(x, y)$ and $F_{XY}(x, y)$.
- Notice that $p_{XY}(1, i) = p_X(1)p_Y(i)$ and $F_{XY}(1, i) = F_X(1)F_Y(i)$.
- In general, if $p_{XY}(x, y) = p_X(x)p_Y(y)$ and $F_{XY}(x, y) = F_X(x)F_Y(y)$ we say X and Y are independent.

> Two random variables, X and Y are independent if the following is true:
> $p_{XY}(x, y) = p_X(x)p_Y(y)$ and $F_{XY}(x, y) = F_X(x)F_Y(y)$

Independence

Two random variables, X and Y are independent if the following is true:
$p_{XY}(x, y) = p_X(x)p_Y(y)$ and $F_{XY}(x, y) = F_X(x)F_Y(y)$

- How does this relate to $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$?
- $A = \{\omega \in \Omega : X(\omega) \leq x\}$ and $B = \{\omega \in \Omega : Y(\omega) \leq y\}$.
- $F_{XY}(x, y) := \mathbb{P}\{\omega \in \Omega : X(w) \leq x \text{ and } Y(\omega) \leq y\} = \mathbb{P}(A \cap B)$.

$E[XY]$

- $E[X] = \sum_x x p_X(x)$ and $E[Y] = \sum_y y p_Y(y)$
- $E[X] = \sum_x \sum_y x p_{XY}(x, y)$ and $E[Y] = \sum_x \sum_y y p_{XY}(x, y)$
- How do we define $E[XY]$?
- You want to search over all values $X \times Y$ can take ($\{1, 2, .., 6\}$) and weight it by the corresponding probabilities.
- $E[XY] = \sum_x \sum_y xy p_{XY}(x, y) = 1.75 = E[X]E[Y]$.

> If X and Y are independent, $E[XY] = E[X]E[Y]$.

Example where X and Y are Dependent

- Now consider rolling a dice.
- $X = \begin{cases} 1 \text{ if outcome is odd} \\ 0 \text{ otherwise .} \end{cases}$ and $Y = \begin{cases} 1 \text{ if outcome is even} \\ 0 \text{ otherwise .} \end{cases}$
- What is $p_X(x), p_Y(y), p_{XY}(x, y)$ and $F_{XY}(x, y)$?
- What is $E[XY]$?

Towards $E[g(X, Y)]$

- What about $E[aX + bY + c]$?

$$
\begin{aligned}
E[aX + bY + c] &= \sum_{x,y}(ax + by + c)p_{XY}(x, y) \\
&= a\sum_{xy} x p_{XY}(x, y) + b\sum_{xy} y p_{XY}(x, y) \\
&+ c\sum_{xy} p_{XY}(x, y) \\
&= aE[X] + bE[Y] + c.
\end{aligned}
$$

- What is $E[aX + bY + c]$ for the running example?

- Along similar lines, one would expect:

$$\boxed{E[g(X, Y)] = \sum_x \sum_y g(x, y)p_{XY}(x, y)}$$

Finding $p_Z(\cdot)$ where $Z = g(X, Y)$.

- Suppose $Z = g(X)$. Then what is $p_Z(z)$?
- $p_Z(z) = \sum_{\{x: g(x)=z\}} p_X(x)$.
- Now suppose $Z = g(X, Y)$ then we have

$$\boxed{p_Z(z) = \sum_{\{x,y: g(x,y)=z\}} p_{XY}(x, y)}$$

Does this need a proof? No if you are an unconscious statistician!

$E[g(X, Y)]$

- $E[X] = \sum_x x p_X(x)$ and $E[Y] = \sum_y y p_Y(y)$
- $E[X] = \sum_{xy} x p_{XY}(x, y)$ and $E[Y] = \sum_{xy} y p_{XY}(x, y)$
- How do we define $E[g(X, Y)]$?
- One way is to define $Z = g(X, Y)$ and find $E[Z] = \sum_z z P_Z(z)$
- Recall $p_Z(z) = \sum_{\{x,y:g(x,y)=z\}} p_{XY}(x, y)$
- This gives us $E[Z] = \sum_z \sum_{\{x,y:g(x,y)=z\}} z p_{XY}(x, y)$.
- This is same as $E[g(X, Y)] = \sum_{\{x,y\}} g(x, y) p_{XY}(x, y)$.

$$E[g(X, Y)] = \sum_{xy} g(x, y) p_{XY}(xy)$$

https://en.wikipedia.org/wiki/Law_of_the_unconscious_statistician

Consistency conditions

- $\sum_{x,y} p_X(x, y) = 1.$
- $F_{XY}(\infty, \infty) = 1.$
- $F_{XY}(-\infty, -\infty) = 0.$
- $F_{XY}(-\infty, \infty) = 0.$
- $F_{XY}(\infty, -\infty) = 0$

Summary

$p_{XY}(x,y) := \mathbb{P}\{\omega \in \Omega : X(w) = x \text{ and } Y(\omega) = y\}.$
$F_{XY}(x,y) := \mathbb{P}\{\omega \in \Omega : X(w) \leq x \text{ and } Y(\omega) \leq y\}.$

The marginal PMF's p_X and p_Y can be obtained from the joint PMF as follows:
$p_X(x) = \sum_y p_{XY}(x,y)$ and $p_Y(y) = \sum_x p_{XY}(x,y)$.

Two random variables, X and Y are independent if the following is true:
$p_{XY}(x,y) = p_X(x)p_Y(y), F_{XY}(x,y) = F_X(x)F_Y(y)$ and $E[XY] = E[X]E[Y]$.

$E[g(X,Y)] = \sum_{xy} g(xy)p_{XY}(xy)$

The rules for more than 2 discrete random variables are similar.

Multiple continuous random variables

- Pick a number uniformly at random from a unit square centered at $(.5, .5)$.
- Random variables X and Y represent the respective x and y coordinate of the point chosen.
- $F_{X,Y}(x, y)$ denotes the probability that the point chosen lies below and to left of point (x, y).
- In this example, $F_{X,Y}(x, y) = xy$.
- Now visualize $F_{X,Y}(x + h, y) - F_{X,Y}(x, y)$. This is the probability that the point chosen lies in the thin strip below y and between x and $x + h$.

Multiple continuous random variables

- Visualize $F_{X,Y}(x+h, y) - F_{X,Y}(x, y)$. This is the probability that the point chosen lies in the thin strip below y and between x and $x+h$.
- $\frac{\partial F_{XY}(x,y)}{\partial x} = lim_{h \to 0} \frac{F_{X,Y}(x+h,y) - F_{X,Y}(x,y)}{h}$.
- This is the rate of change of the joint CDF $F_{XY}(x, y)$ in the x direction.

Multiple continuous random variables

- $\frac{\partial F_{XY}(x,y}{\partial y} = \lim_{h \to 0} \frac{F_{X,Y}(x,y+h) - F_{X,Y}(x,y)}{h}$ denotes the rate of change of the joint CDF in the y direction.
- $f_{X,Y}(x,y) := \frac{\partial^2 F_{XY}(x,y)}{\partial x \partial y}$ represents the joint probability density function.
- $f_{X,Y}(x,y)dxdy$ denotes the probability that (X, Y) are in a rectangle of area $dxdy$ around (x, y).
- In this example, $f_{X,Y}(x,y) = 1$.
- $F_{XY}(x,y) := \int_{-\infty}^{x} \int_{-\infty}^{y} f_{XY}(s,t)dsdt$.

Summary for Continuous random variable

- $f_{XY}(x,y)$ denotes the joint pdf for X and Y.
- $F_{XY}(x,y) := \int_{-\infty}^{x}\int_{-\infty}^{y} f_{XY}(s,t)dsdt$. $f_{X,Y}(x,y) := \frac{\partial^2 F_{XY}(x,y)}{\partial x \partial y}$.

> The marginal pdf's f_X and f_Y can be obtained from the joint PDF as follows:
> $f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)dy$ and $f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)dx$

> Two random variables, X and Y are independent if the following is true:
> $f_{XY}(x,y) = f_X(x)f_Y(y), F_{XY}(x,y) = F_X(x)F_Y(y)$ and $E[XY] = E[X]E[Y]$.

> $E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f_{XY}(x,y)dxdy$

- Rules similar for more than 2 random variables.