PROBABILITY THEORY
Session 2

RANDOM VARIABLES

Topics:

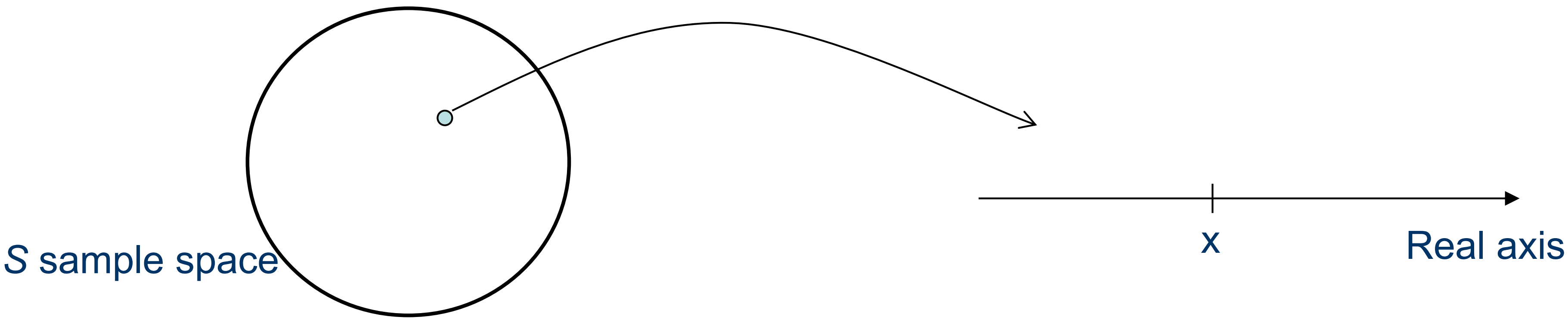
- Discrete and continuous random variables.
- Probability mass function and probability density function.
- Cumulative distribution function.
- Jointly distributed random variables.
- Independent random variables.
- Expectation. Variance. Covariance.

What should we learn today?

- What is a random variable?
- What types of random variables exists – how are they different and similar?
- How to work with distributions:
  - Cumulative density function
  - Probability mass function
  - Probability density function
- How to handle jointly distributed multiple random variables?
- How to define and work with expectation, variance and covariance

Random variable

- Definition. A r.v. *X* is a function that assigns a real number to each outcome of a random experiment.



Examples

- A r.v. X is a lifetime of a system component.
- Time until a next job arrives at a server
- Number of trees on a square meter
- Coin flipping

Example

- A coin is flipped 3 times. A r.v. X is a number of heads in 3 trials.

Example: coin flipping

Possible outcomes:

HHH	3
HHT	2
HTH	2
THH	2
HTT	1
THT	1
TTH	1
TTT	0

$P(X=3)=\frac{1}{8}$

$P(X=2)=\frac{3}{8}$

$P(X=1)=\frac{3}{8}$

$P(X=0)=\frac{1}{8}$

Example: coin flipping

Possible outcomes:

HHH	3
HHT	2
HTH	2
THH	2
HTT	1
THT	1
TTH	1
TTT	0

$$P(X=3)=\frac{1}{8}$$

$$P(X=2)=\frac{3}{8}$$

$$P(X=1)=\frac{3}{8}$$

$$P(X=0)=\frac{1}{8}$$

Total number of outcomes =

$$=2^3=8$$

\# of outcomes when # of H is i

$$\binom{3}{i}=\frac{3!}{(3-i)!\,i!}=\begin{cases}1, & i=3\\ 3, & i=2\\ 3, & i=1\\ 1, & i=0\end{cases}$$

$$P(X\leq 2)=P(X=0)+P(X=1)+P(X=2)=\frac{7}{8}$$

How to decribe behaviour of a r.v.?

Discrete r.v.

- Definition. A r.v. whose set of possible values is a sequence is said to be discrete.
- For a discrete r.v. we define probability mass function

$$\boxed{p(a) = P(X = a)}$$

$$\boxed{\begin{array}{l} p(x_i) > 0, \quad i = 1, 2, \ldots \\ p(x) = 0, \quad otherwise \end{array}}$$

$$\boxed{\sum_{i=1}^{\infty} p(x_i) = 1}$$

Example coin flipping

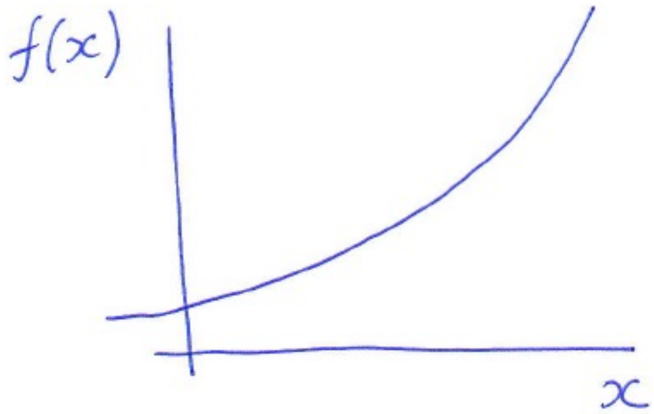
Continuous r.v.

- Definition. X is a continuous r.v. if there exists a nonnegative function f(x), defined for all x having the property that for any set B of real numbers

$$\boxed{P(X \in B) = \int_B f(x)dx}$$

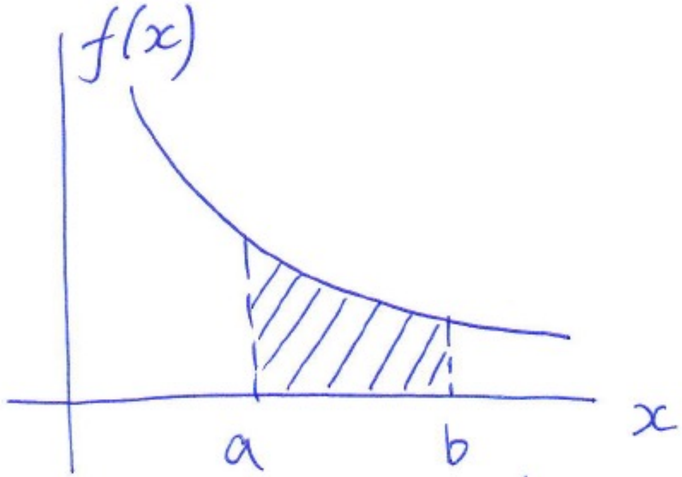
- Function f(x) is called the probability density function of X
- Additionally, f(x) should satisfy $\boxed{\int_{-\infty}^{\infty} f(x)dx = 1}$
- Different probability statements can be expressed using pdf
- If we let a=b, then $\boxed{P(a \leq X \leq b) = \int_a^b f(x)dx}$ $\boxed{P(X = a) = \int_a^a f(x)dx = 0}$
- Probability that a continuous r.v. will assume any particular value is zero

Continuous r.v.



not pdf!

$$\int_{-\infty}^{+\infty} f(x)\,dx = 1$$



$$P\{a < X \le b\} = \int_a^b f(x)\,dx$$

$$P(X = a) = \int_a^a f(x)\,dx = 0$$

Example

The pdf of the samples of the amplitide of speech waveforms is found to decay exponentially at rate alpha

$$f(x) = ce^{-\alpha|x|}$$

Find constant c and find probability $P(|X| < v)$

Example speech amplitude

- Use normalization condition:

$$1 = \int_{-\infty}^{+\infty} c e^{-\alpha|x|} dx = 2c \int_{0}^{\infty} e^{-\alpha x} dx =$$

$$= 2c \left. \frac{e^{-\alpha x}}{-\alpha} \right|_{0}^{\infty} = 2c\left(0 + \frac{1}{\alpha}\right) = \frac{2c}{\alpha}$$

$$\Rightarrow \quad c = \frac{\alpha}{2}$$

- Find probability

$$P\{|X| < v\} = \frac{\alpha}{2} \int_{-v}^{v} e^{-\alpha|x|} dx = 2 \cdot \frac{\alpha}{2} \int_{0}^{v} e^{-\alpha x} dx =$$

$$= \alpha \left. \frac{e^{-\alpha x}}{-\alpha} \right|_{0}^{v} = \left[1 - e^{-\alpha v}\right]$$

Cumulative distribution function

- The cdf of a r.v. *X* is defined for any real number *x* as the probability of the event *{ X≤x }*

$$\boxed{F(x) = P(X \leq x)}$$

- *F* is a function of *x*.
- All probability questions about *X* can be answered in terms of its distribution function
- Example: how to compute *P(a<X ≤b)* ?

$$\{X \leq b\} = \{X \leq a\} \bigcup \{a < X \leq b\}$$

$$P\{X \leq b\} = P\{X \leq a\} + P\{a < X \leq b\}$$

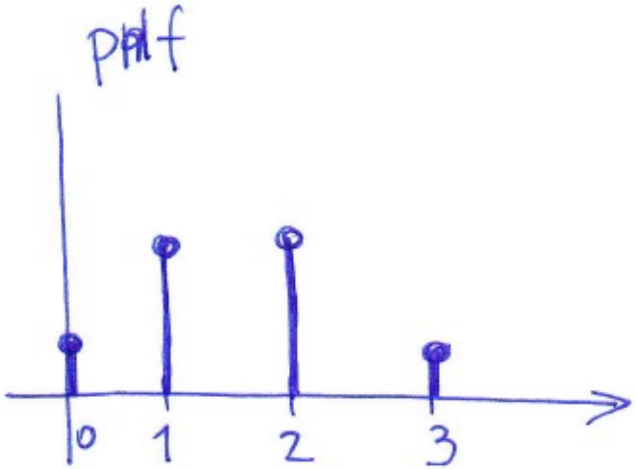
$$\boxed{P\{a < X \leq b\} = F(b) - F(a)}$$

Cdf for discrete r.v.

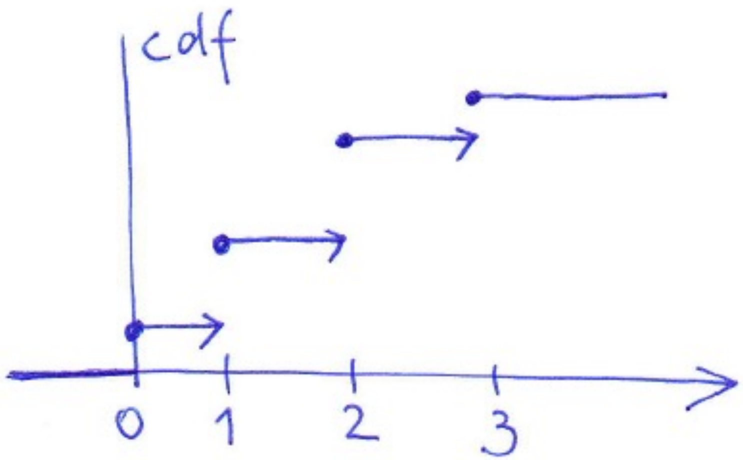
- The cdf is a step-function and can be expressed as

$$\boxed{F(a) = \sum_{all\ x \leq a} p(x)}$$

Example coin flipping



$$\sum_i p(i) = 1$$



$$F(a) = \sum_{\text{all } x \leq a} p(a)$$

Cdf for continuous r.v

- The relationship between cdf and pdf is expressed by

$$\boxed{F(a) = P(X \in (-\infty, a]) = \int_{-\infty}^{a} f(x)dx}$$

$$\boxed{\frac{d}{da}F(a) = f(a)}$$

Properties of cdf

1. $0 \le F(x) \le 1$
2. $\lim_{x \to \infty} F(x) = 1$
3. $\lim_{x \to -\infty} F(x) = 0$
4. $F(x)$ is a nondecreasing function: $F(a) \le F(b)$ if $a < b$
5. $F(x)$ is a continuous from the right:

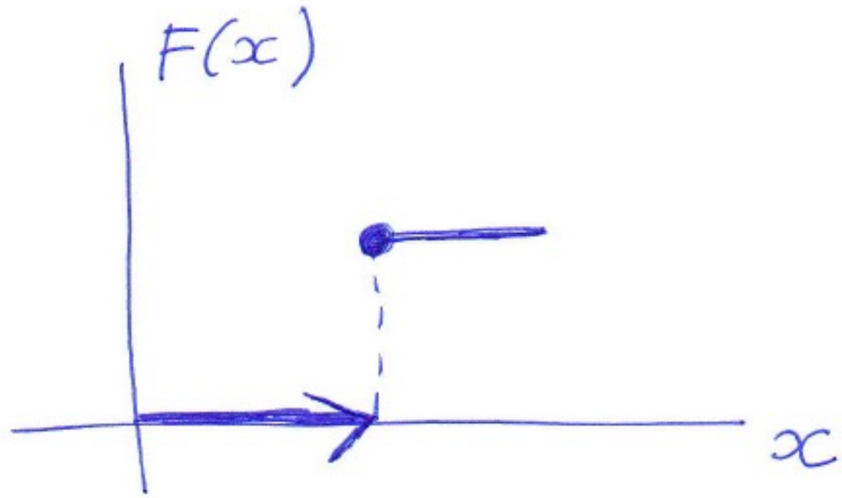
$$F(b) = \lim_{h \to 0} F(b+h) = F(b+)$$

6. Probability that a r.v. X takes on a specific value b is equal to the jump (step) of cdf at the point b:

$$P(X = b) = F(b+) - F(b-)$$

Property no 5

Continuous from the right:

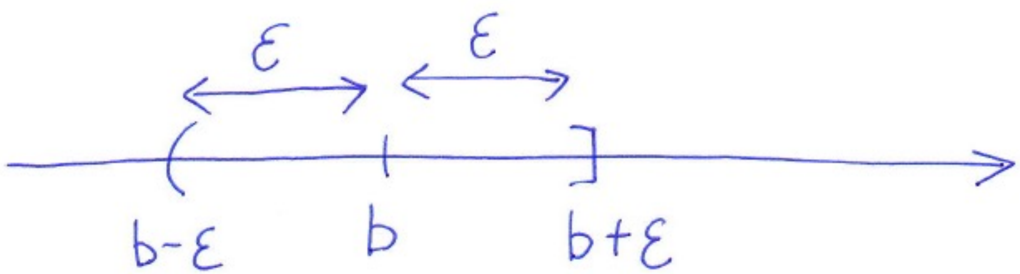


Property no 6

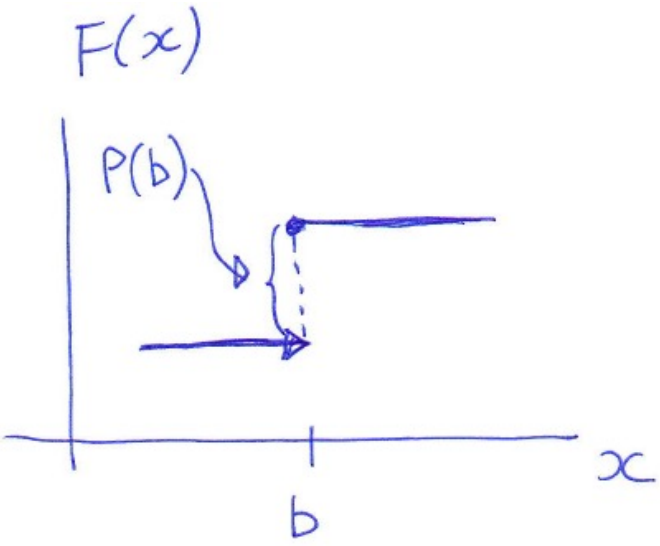
We know $P\{a < X \leq b\} = F(b) - F(a)$

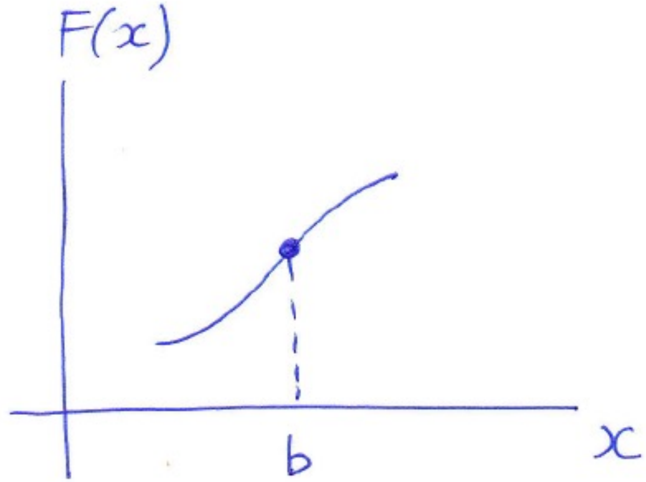
$$P\{X = b\} = \lim_{\varepsilon \to 0} \{b - \varepsilon < X \leq b + \varepsilon\} =$$

$$= \lim_{\varepsilon \to 0} [F(b+\varepsilon) - F(b-\varepsilon)] = F(b+) - F(b-)$$



Property no 6





discrete r.v.

$P(b) =$ "jump"

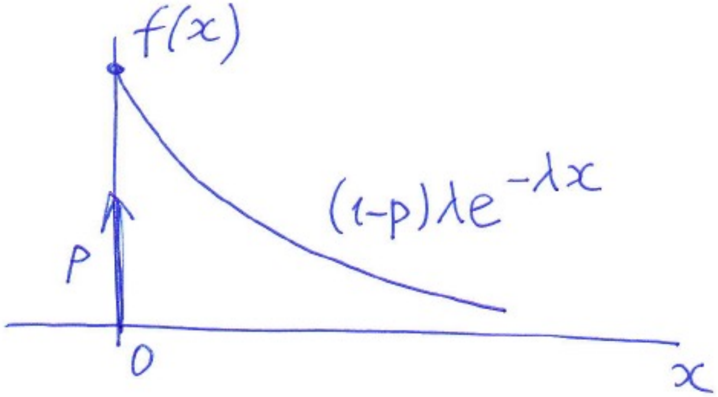
continuous r.v.

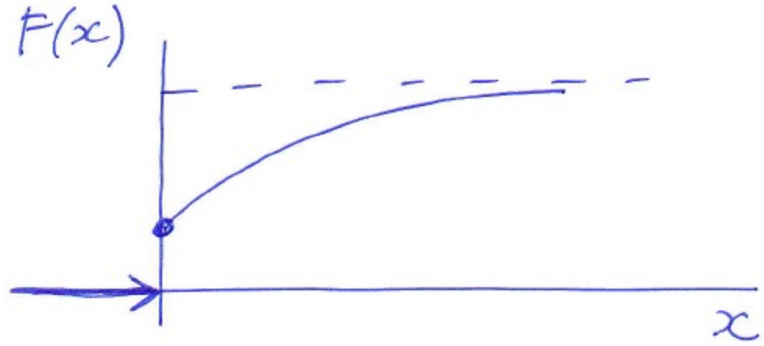
$P(b) = 0$

Example of a mixed r.v.

- The delay (= waiting time in a queue) for a packet transmission is zero if the queue is empty, and if the queue is not empty, the delay is an exponentially distributed r.v. with cdf $F(x) = 1 - e^{-\lambda x}$
- The probability that the queue is empty is ***p*** and busy ***1-p***.
- Cdf of the delay X:

$$F(x) = P(X \le x) = P(X \le x \mid \text{idle}) \cdot p + P(X \le x \mid \text{busy}) \cdot (1-p)$$
$$= p + (1-p)(1-e^{-\lambda x})$$





Types of r.v.

Discrete r.v.	Continuous r.v.	Mixed type
Pmf p(x) $$P\{a < X < b\} = \sum_{x_i \in (a,b)} p(x_i)$$	Pdf f(x) $$P(a \le X \le b) = \int_a^b f(x)dx$$	

Multiple r.v.

- So far, we were speaking about calculation of probabilities of events involving a single r.v. in isolation. Now we will look at the techniques for probability calculations of events that involve the joint behavior of two or more r.v.
- Example: height, weight and age of a person from a group

Joint cdf

- To spesify the relationship between two r.v., we define the joint cumulative probability distribution function of X and Y

$$\boxed{F(X,Y) = P(X \le x, Y \le y)}$$

- A knowledge of the joint cdf enables us to calculate the distribution function of r.v. X:

$$F_X(x) = P(X \le x) = P(X \le x, Y < \infty) = F(x, \infty)$$

Joint pmf for discrete r.vs.

- If X and Y are discrete r.vs., we define joint probability mass function:

$$\boxed{p(x_i, y_i) = P(X = x_i, Y = y_i)}$$

- The individual mass functions are easily obtained from the joint pmf:

$$\{X = x_i\} = \bigcup_j \{X = x_i, Y = y_i\}$$

$$P\{X = x_i\} = P\bigcup_j \{X = x_i, Y = y_i\} = \sum_j P\{X = x_i, Y = y_i\} = \sum_j p(x_i, y_j)$$

- The joint probabilities can be presented in tabular form. Because the individual probabilities appear in the margin of the table, they are often called marginal probabilities.

Joint pdf for continous r.vs.

- We say that X and Y are jointly continous, if there exist a function f(x,y) defined for all real x and y, having the property that for every set C in the 2-dimentional plane

$$\boxed{P\{(X,Y) \in C\} = \int\int_{(x,y)\in C} f(x,y)dxdy}$$

- f(x,y) is called joint probability density function

$$P\{X \in [a,b], Y \in [c,d]\} = \int_c^d \int_a^b f(x,y)dxdy$$

- The marginal pdfs are obtained by integrating out the variables that are not of interest.

$$\boxed{f_X = \int_{-\infty}^{\infty} f(x,y)dy} \quad \boxed{f_Y = \int_{-\infty}^{\infty} f(x,y)dx}$$

Question

- As we have seen, specifying the joint probability mass function or probability density function determines the individual distribution functions.
- Is reverse true? If I know individual mass functions, can I determine the joint mass function?

Example: transmission system

R \ Tr	0	1	
0	1/4	1/4	1/2
1	1/4	1/4	1/2
	1/2	1/2	

R \ Tr	0	1	
0	1/2	0	1/2
1	0	1/2	1/2
	1/2	1/2	

Independent r.v.

- Definition. X and Y are independent, if for any two sets of real numbers A and B

$$\boxed{P(X \in A, Y \in B) = P(X \in A)P(Y \in B)}$$

- In terms of joint cumulative distribution function:

$$\boxed{F(a,b) = F_X(a)F_Y(b)}$$

- In terms of pmf (discrete r.v.) and pdf (continuous r.v.)

$$\boxed{p(x,y) = p_X(x)p_Y(y)} \qquad \boxed{f(x,y) = f_X(x)f_Y(y)}$$

- Basically, X and Y are independent, if knowing the value of one does not change the distribution of another

Multiple r.v.

- Let $X_1, \ldots, X_n$ be the jointly distributed random variables.
- Definitions for pmf, pdf and cdf can be generalized to a case of n r.vs. For example,

- the joint cumulative distribution function is defined as

$$\boxed{F(x_1, \ldots, x_n) = P(X_1 \le x_1, \ldots, X_n \le x_n)}$$

- Example. A computer system receives messages over three communication lines. Let X_i be the number of messages received on line i in one hour. The joint pmf is given by

$$p(x_1, x_2, x_3) = (1 - a_1)(1 - a_2)(1 - a_3)a_1^{x_1} a_2^{x_2} a_3^{x_3}$$

Find individual pmfs

Example 3 communication lines

$$p_{X_3}(x_3) = \sum_{i=1}^{\infty} \sum_{k=1}^{\infty} (1-a_1)(1-a_2)(1-a_3)\, a_1^{i}\, a_2^{k}\, a_3^{x_3} =$$

$$= (1-a_1)(1-a_2)(1-a_3)\, a_3^{x_3} \underbrace{\sum_{i=1}^{\infty} a_1^{i}}_{\frac{1}{1-a_1}} \underbrace{\sum_{k=1}^{\infty} a_2^{k}}_{\frac{1}{1-a_2}} =$$

$$= (1-a_3)\, a_3^{x_3}$$

- Q: are lines independent?

Conditional distributions

- The relationship between two random variables can often be clarified by consideration of the conditional distribution of one given the value of the other.
- The conditional pmf of X given that Y=y is defined by

$$p_{X|Y}(x|y) = \frac{p(x,y)}{p_Y(y)}$$

- If X and Y have a joint pdf, then the conditional pdf of X given that Y=y is defined as

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$$

Example conditional distributions

- Joint pmf of X and Y is given:

Y \ X	0	1	
0	0.4	0.1	0.5
1	0.2	0.3	0.5
	0.6	0.4	

- Calculate the conditional pmf of X given that Y=1:

$$P\{X=0 \mid Y=1\} = \frac{p(0,1)}{P\{Y=1\}} = \frac{0.2}{0.5} = 0.4$$

$$P\{X=1 \mid Y=1\} = \frac{p(1,1)}{P\{Y=1\}} = \frac{0.3}{0.5} = 0.6$$