NMAI059 Probability and statistics 1
Class 5

Robert Šámal

Overview

Random vectors

Conditional distribution

Continuous random variables

Basic description of random vectors

- X, Y – random variables on the same probability space $(\Omega, \mathcal{F}, P)$.
- We wish to treat (X, Y) as one object – a random vector.
- How to do that?
- Example: we roll twice a 4-sided dice, $X =$ first outcome, $Y =$ second one.

$X, Y \sim Bin(\frac{1}{2})$

$X = Y$ (toss one coin)

$X=Y=0$	$X=Y=1$

Ω

X, Y independent

toss two coins

00	01
10	11

$X = 1 - Y$ ... toss one coin

$X = \#$ of Heads

$Y = \#$ of Tails

$X=0$, $Y=1$	$X=1$, $Y=0$

Joint distribution

$P : \mathcal{F} \to [0,1]$

Definition

For a discrete r.v. X, Y on a probability space $(\Omega, \mathcal{F}, P)$ we define their joint PMF $p_{X,Y} : \mathbb{R}^2 \to [0,1]$ by a formula

$= P(X=x \,\&\, Y=y)$

$$p_{X,Y}(x,y) = P(\{\omega \in \Omega : X(\omega) = x \& Y(\omega) = y\}).$$

For this we need that for each $x, y \in \mathbb{R}$ we have $\{\omega \in \Omega : X(\omega) = x \& Y(\omega) = y\} \in \mathcal{F}$, otherwise we do not consider (X, Y) as a random vector.

- We can define it also for more than two r.v.'s $p_{X_1,\dots,X_n}(x_1, \dots, x_n)$. $= P(X_1 = x_1 \,\&\, \dots \,\&\, X_n = x_n)$

Marginal distribution

- Given $p_{X,Y}$, how to find the distribution of each of the coordinates, that is p_X and p_Y?

$p_X(1) = P(X=1) = P(X=1 \,\&\, Y=1) + P(X=1 \,\&\, Y=2) + \dots + P(X=1 \,\&\, Y=4)$

X, Y uniform on $\{1,2,3,4\}$

p_X :

1	2	3	4
$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

p_Y

$X \backslash Y$	1	2	3	4	Σ
1	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{4}{16} = \frac{1}{4}$
2	$\frac{1}{16}$	$\frac{1}{16}$	...	-	$\frac{1}{4}$
3	-	-	-		$\frac{1}{4}$
4	-	-	- -	-	$\frac{1}{4}$

Y

p_X

$X \backslash Y$	1	2	3	4	Σ
1	$\frac{1}{4}$	0	0	0	$\frac{1}{4}$
2	0	$\frac{1}{4}$	0	0	$\frac{1}{4}$
3	0	0	$\frac{1}{4}$	0	$\frac{1}{4}$
4	0	0	0	$\frac{1}{4}$	$\frac{1}{4}$

$\Sigma = 1$

$P(X=1 \,\&\, Y=2) = 0$

independent

p_X, p_Y are the same
$p_{X,Y}$ is NOT

$X = Y$

Theorem X, Y discr. r.v.

$$P_X(x) \overset{df}{=} P(X=x) = \sum_{y \in \mathrm{Im}(Y)} P(X=x \,\&\, Y=y) =: \sum_{y \in \mathrm{Im}(Y)} P_{X,Y}(x,y) \quad (df)$$

$$P_Y(y) = P(Y=y) = \sum_{x \in \mathrm{Im}(X)} P(X=x \,\&\, Y=y) = \sum_{x \in \mathrm{Im}(X)} P_{X,Y}(x,y)$$

Proof

$$\{\omega : X(\omega)=x\} = \bigcup_{y \in \mathrm{Im}(Y)} \{\omega : X(\omega)=x \,\&\, Y(\omega)=y\}$$

$$P(\quad) = \sum P(\text{---})$$

Independence of r.v.'s

Definition

Discrete r.v.'s X, Y are independent if for every $x, y \in \mathbb{R}$ the events $\{X = x\}$ a $\{Y = y\}$ are independent. That happens if and only if

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

prob. of intersection $\{X=x\} \cap \{Y=y\}$

$$P_{X,Y}(x,y) = P_X(x) \cdot P_Y(y)$$

If X, Y are independent, then $P_{X,Y}$ can be recovered from P_X, P_Y

Example: Multinomial distribution

$p_1+\dots+p_6=1$, $n\in\mathbb{N}$

- On a die we roll i with probability p_i for $i=1,\dots,6$. We roll the die n-times and let X_i be the number of rolls when i came up.

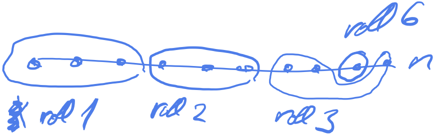
$(X_1+\dots+X_6=n)$

$$P_{X_1,\dots,X_6}(k_1,\dots,k_6)=P(X_1=k_1\ \&\ \dots\ \&\ X_6=k_6)=\binom{n}{k_1,\dots,k_6}p_1^{k_1}\cdots p_6^{k_6}$$

$\binom{n}{k_1,\dots,k_6} = \frac{n!}{k_1!\cdots k_6!}$

→ # of ways to part. set $\{1,\dots,n\}$ to a set of size k_1 + … k_2 + … k_6 —

prob. we roll 1 at the k_1 places we chose for 1



$6 \rightsquigarrow 2$ $\quad p_1+p_2=1$

$\binom{n}{k_1,k_2} = \binom{n}{k_1}$ $\quad X_1 \sim Bin(n,p_1)$

$X_2 = n - X_1$

NOT INDEP.: $P_{X_1\dots X_6}(0,\dots,0)=0\neq\prod_{i=1}^{6}P_{X_i}(0)$

ALWAYS $\boxed{X_i \sim Bin(n,p_i)}$

Coupling – nontrivial use of joint distributions

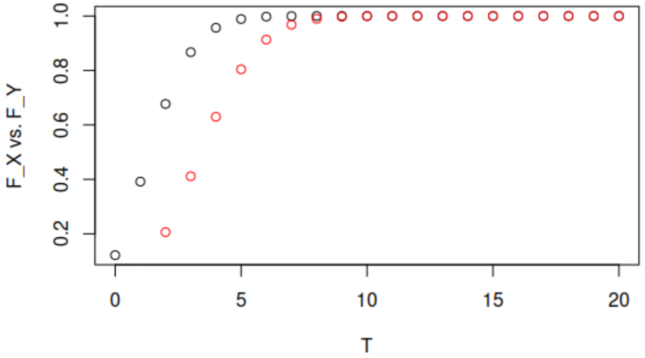
- $X \sim Bin(n,p)$ and $Y \sim Bin(n,q)$ for $p < q$
- What can be said about F_X and F_Y?
- $\sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i}$ is an increasing function of p – but why?

$F_X \le F_Y$?

$F_X \ge F_Y$?

$F_X(k) =$

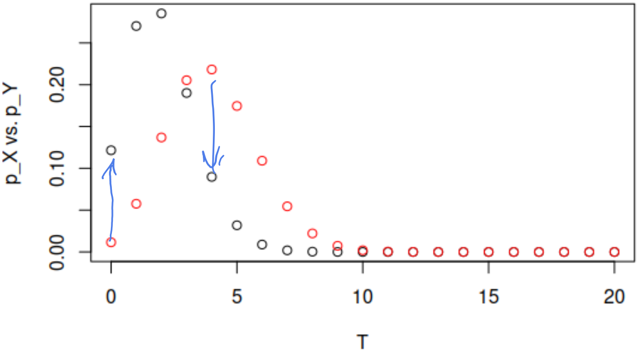
$X \sim Bin(20, 0.1)$ $Y \sim Bin(20, 0.2)$ CDF



Coupling – nontrivial use of joint distributions

- $X \sim Bin(n,p)$ and $Y \sim Bin(n,q)$ for $p < q$
- What can be said about F_X and F_Y?
- $\sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i}$ is an increasing function of p – but why?

PICF



Coupling

$X \sim Bin(n,p)$

$P(X \le k) \ge P(Y \le k)$

$F_X(k) \ge F_Y(k)$

- $X = \sum_{i=1}^{n} X_i$, where $X_1, \dots, X_n$ are independent $\sim Bern(p)$
- $Y = \sum_{i=1}^{n} Y_i$, where $Y_1, \dots, Y_n$ are independent $\sim Bern(q)$
- Joint distribution of X and Y is not determined, it can be arbitrary.
- We make it so that X and Y are not independent, more so, $\boxed{\text{always } X \le Y}$
- It suffices to define $Y_i =$

$\boxed{\text{always } X_i \le Y_i}$

prob. p: if $X_i = 1$ then $Y_i := 1$

prob. $1-p$: if $X_i = 0$ then: $\frac{q-p}{1-p}$: $Y_i := 1$; $\frac{1-q}{1-p}$: $Y_i := 0$

$Y \sim Bin(n,q)$
$\Uparrow$
$Y_1 \dots Y_n$ ind.
$\Uparrow$
Y_i is determ. only by X_i

$P(Y_i = 1) = P(X_i=1)P(Y_i=1 \mid X_i=1) + P(X_i=0)P(Y_i=1 \mid X_i=0)$

$= p \cdot 1 + (1-p) \cdot \frac{q-p}{1-p} = q$

$P(Y_i=0) = P(X_i=0)P(Y_i=0 \mid X_i=0) = (1-p)\frac{1-q}{1-p} = 1-q$

Product of independent r.v.'s

(X, Y NOT INDEP. ⟹ MAY BE FALSE

$X = Y$ $\quad \mathbb{E}(X^2) \overset{?}{\neq} (\mathbb{E}(X))^2$

$(\mathrm{var}(X) = \mathbb{E}(X^2) - (\mathbb{E}X)^2)$

Theorem

For independent discrete r.v.'s X, Y we have

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y).$$

Proof $\quad \mathbb{E}(X \cdot Y) \overset{?}{=} \sum_{\substack{x \in \mathrm{Im}(X) \\ y \in \mathrm{Im}(Y)}} xy \, P(X = x \,\&\, Y = y)$

independence

$$= \sum xy \, P(X = x) \cdot P(Y = y)$$

basic algebra

$$= \sum_{x \in \mathrm{Im}(X)} x \cdot P(X = x) \cdot \sum_{y \in \mathrm{Im}(Y)} y \cdot P(Y = y)$$

$(a+b)(c+d)$
$= ac + ad + bc + bd$

def.

$$= \mathbb{E}(X) \cdot \mathbb{E}(Y).$$

Function of a random vector

Theorem

(2D cotas)

Suppose X, Y are discrete r.v.'s on $(\Omega, \mathcal{F}, P)$, let $g : \mathbb{R}^2 \to \mathbb{R}$ be a function.

- *Then $Z = g(X, Y)$ is a* discrete *r.v. on $(\Omega, \mathcal{F}, P)$*
- *and it satisfies*

$$\mathbb{E}(g(X,Y)) = \sum_{x \in ImX} \sum_{y \in ImY} g(x,y) P(X = x, Y = y),$$

$g(x,y) = xy$

whenever the sum is defined.

$\sum_{x} \Big(ax \underbrace{\sum_{y} P(X=x, Y=y)}_{P(X=x)} \Big) = a \sum_{x} x \cdot P(X=x)$

Theorem (Linearity of expectation)

discrete

For X, Y r.v.'s (independence is not needed!) and $a, b \in \mathbb{R}$ we have

$$\boxed{\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y).}$$

Proof

$g(x,y) = ax + by$

$E(aX + bY) = E\, g(X,Y) = \sum_{x,y} (ax + by) P(X=x, Y=y) = \sum_{x,y} ax\, P(X=x, Y=y) + \sum_{x,y} by\, P(X=x, Y=y)$

$Z := g(X, Y)$: For $\underline{\omega \in \Omega}$ $\underline{Z(\omega) = g(X(\omega), Y(\omega))}$

$\hookrightarrow$ is a disc. r.v.

$\forall z \in \mathbb{R}$ $Z^{-1}(z) = \underline{\{\omega : Z(\omega) = z\}} \in \mathcal{F}$

$$= \bigcup_{\circledast} \{\omega : \underbrace{\underline{X(\omega) = x} \;\&\; \underline{Y(\omega) = y}}_{\in \mathcal{F}}\} \in \mathcal{F}$$

$$\circledast \left\{ \begin{array}{l} x \in \mathrm{Im}(X) \\ y \in \mathrm{Im}(Y) \\ \boxed{g(x,y) = z} \end{array} \right\}$$

$\leftarrow$ countable union of disj. sets

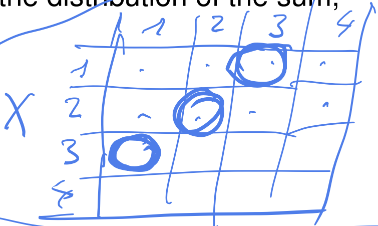
$$P(Z = z) = \sum_{\circledast} P(X = x, Y = y)$$

$$EZ = \sum_z z \cdot P(Z = z) = \sum_z \cdot \sum_{\circledast} z \cdot P(X = x, Y = y) = \sum_{\substack{x \in \mathrm{Im}(X) \\ y \in \mathrm{Im}(Y)}} g(x,y) P(\ldots)$$

$z = g(x,y)$

Sum of independent r.v.'s

- Given $p_{X,Y}$, how to find the distribution of the sum, $Z = X + Y$?



$g(x,y) = X + Y$

$(EZ = EX + EY)$

$\{Z = z\} = \bigcup_{\substack{x,y \\ x+y=z}} \{X = x, Y = y\}$

$P(X + Y = 4)$

$= P(X = 3 \,\&\, Y = 1) + P(X = 2, Y = 2)$

$+ P(X = 1, Y = 3)$

Sum of r.v.'s – convolution

$$\boxed{P_Z(z) = \sum_x P_X(x) P_Y(z-x)}$$

Theorem (Convolution formula)

Let X, Y be discrete random variables. Then their sum $Z = X + Y$ has PMF given by

$$P(Z = z) = \sum_{x \in Im(X)} P(X = x, Y = z - x).$$

If we further assume that X, Y are independent, then

$$P(Z = z) = \sum_{x \in Im(X)} P(X = x) P(Y = z - x).$$

convolution

Proof: Axioms for disj. union +

$$\{Z = z\} = \bigcup_{x \in Im(X)} \{X = x, Y = y\}$$

Example of a convolution

$\underbrace{\circ\,\circ\,\circ \;\cdots\; \circ}_{m \text{ attempts}}$ $\quad \underbrace{\square\,\square\,\square\;\square}_{n,\ Y}$

X successes

$X \sim \mathrm{Bin}(m,p)$

$Y \sim \mathrm{Bin}(n,p)$ $\longrightarrow$ $\underbrace{X+Y}_{Z} \sim \mathrm{Bin}(m+n,p)$

X, Y independent

$$P(Z=z) = \sum_k P(X=k \;\&\; Y=z-k)$$

$$= \sum_k P(X=k) \cdot P(Y=z-k)$$

$$= \sum_{k=0}^{m} \binom{m}{k} p^k (1-p)^{m-k} \binom{n}{z-k} p^{z-k} (1-p)^{n-(z-k)}$$

$$= \left[\sum_{k=0}^{m} \binom{m}{k}\binom{n}{z-k}\right] \cdot \underbrace{p^z (1-p)^{m+n-z}}_{\text{indep. of } k} = \binom{m+n}{z} p^z (1-p)^{m+n-z}$$

$\mathrm{Bin}(m+n,p)$

Overview

Random vectors

Conditional distribution

Continuous random variables

Conditional PMF

x	1	2	3	4
$p_X(x)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

x	1	2	3	4
$p_{X\mid A}$	0	$\frac{1}{2}$	0	$\frac{1}{2}$

X, Y – discrete random variables on $(\Omega, \mathcal{F}, P)$, $A \in \mathcal{F}$

- $p_{X|A}(x) := P(X = x \mid A)$ $= P(X=x \,\&\, A)/P(A)$
  example: X is outcome of a roll of a die, $A =$ we got an even number
- $p_{X|Y}(x|y) := P(X = x \mid Y = y)$ example: X, Z is an outcome of two independent die rolls, $Y = X + Z$.

$$p_{X|Y}(6|10) =$$

- $\boxed{p_{X|Y} \text{ from } p_{X,Y}:}$

$$p_{X|Y}(x|y) = \frac{P(X=x \,\&\, Y=y)}{P(Y=y)} = \frac{p_{X,Y}(x,y)}{p_Y(y)} = \frac{p_{X,Y}(x,y)}{\sum_{x'} p_{X,Y}(x',y)}$$

Joint vs. conditional PMF

$p_{X,Y}$	...	10	11	12
1				
2				
3				
4				
5				
6				

$p_{X\|Y}$	...	10	11	12
1				
2				
3				
4				
5				
6				

Overview

Random vectors

Conditional distribution

Continuous random variables

General random variable

Definition

Random variable on $(\Omega, \mathcal{F}, P)$ *is a mapping* $X : \Omega \to \mathbb{R}$*, such that for each* $x \in \mathbb{R}$

$$\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}.$$

- discrete r.v. is a r.v.

CDF

Definition

Cumulative distribution function, CDF of a r.v. X is a function

$$F_X(x) := P(X \le x) = P(\{\omega \in \Omega : X(\omega) \le x).$$

- F_X is a nondecreasing function
- $\lim_{x \to -\infty} F_X(x) = 0$
- $\lim_{x \to +\infty} F_X(x) = 1$
- F_X is right-continuous

CDF examples

Quantile function

For a r.v. X we define its *quantile function* $Q_X : [0, 1] \to \mathbb{R}$ by

$$Q_X(p) := \min \{x \in \mathbb{R} : p \leq F_X(x)\}$$

- If F_X is continuous, then $Q_X = F_X^{-1}$.
- $Q_X(1/2) =$ median (watch out if F_X is not strictly increasing!)
- $Q_X(10/100) =$ tenth percentile, etc.

Continuous random variable

Definition

R.v. X is called continuous, if there is nonnegative real function f_X such that

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(t)dt.$$

(Sometimes such X is said to be absolutely continuous.)
Function f_X is called the probability density function, pdf of X.

Using density

Theorem

Let X be a continuous r.v. with density f_X. Then

1. $P(X = x) = 0$ *for every* $x \in \mathbb{R}$.
2. $P(a \leq X \leq b) = \int_a^b f_X(t)dt$ *for every* $a, b \in \mathbb{R}$.

Uniform distribution

- R.v. X has a uniform distribution on $[a, b]$, we write $X \sim U(a, b)$, if $f_X(x) = 1/(b - a)$ for $x \in [a, b]$ and $f_X(x) = 0$ otherwise.

Universality of a uniform distribution

Theorem

Let F be a function "of CDF-type": nondecreasing right-continuous function with $\lim_{x \to -\infty} F(x) = 0$ a $\lim_{x \to +\infty} F(x) = 1$. Let Q be the corresponding quantile function.

1. *Let $U \sim U(0,1)$ and $X = Q(U)$. Then X has CDF F.*
2. *Let X be a r.v. with CDF $F_X = F$, suppose F is increasing. Then $F(X) \sim U(0,1)$.*