Jointly distributed random variables

Samy Tindel

Purdue University

Probability Theory 1 - MA 519

Mostly taken from *A first course in probability* by S. Ross

Outline

Outline

Joint cdf

Definition 1.

Let

- X, Y random variables
- $a, b \in \mathbb{R}$

The joint cdf describes the joint distribution of (X, Y):

$$F(a, b) = \mathbf{P}\left(X \leq a,\ Y \leq b\right)$$

Values of interest in terms of the cdf

Proposition 2.

Let

- X, Y random variables
- F the joint cdf of X, Y

Then the marginals cdf's of X and Y are given by

$$F_X(a) = F(a, \infty), \qquad F_Y(b) = F(\infty, b)$$

We also have

$$\mathbf{P}\left(a_1 < X \leq a_2,\, b_1 < Y \leq b_2\right) = F(a_2, b_2) - F(a_2, b_1) - F(a_1, b_2) + F(a_1, b_1)$$

Discrete case: joint pmf

Definition 3.

Consider the following situation:

- X, Y discrete random variables
- X takes values in E_1, Y takes values in E_2
- $x \in E_1$ and $y \in E_2$

The joint pmf p describes the joint distribution of (X, Y):

$$p(x, y) = \mathbf{P}(X = x,\ Y = y)$$

Values of interest in terms of the pmf

Proposition 4.

Let

- X, Y random variables
- p the joint pmf of X, Y

Then the marginals pmf's of X and Y are given by

$$p_X(a) = \sum_{b \in E_2} p(a, b), \qquad p_Y(b) = \sum_{a \in E_1} p(a, b)$$

If $a_1 < a_2$ and $b_1 < b_2$, we also have

$$\mathbf{P}\left(a_1 < X \leq a_2,\, b_1 < Y \leq b_2\right) = \sum_{a_1 < i_1 \leq a_2,\, b_1 < i_2 \leq b_2} p\left(i_1, i_2\right)$$

Example: tossing 3 coins (1)

Experiment:
Tossing a coin 3 times

Events: We consider

$$A = \text{"At most one Head"}$$
$$B = \text{"At least one Head and one Tail"}$$

Random variables: Set

$$X_1 = \mathbf{1}_A, \qquad X_2 = \mathbf{1}_B, \qquad X = (X_1, X_2)$$

Example: tossing 3 coins (2)

Model: We take

- $S = \{h, t\}^3$
- $\mathbf{P}(\{s\}) = \frac{1}{8}$ for all $s \in S$

Description of $X = (X_1, X_2)$:

s	$X(s)$	s	$X(s)$
(t,t,t)	$(1,0)$	(h,t,t)	$(1,1)$
(t,t,h)	$(1,1)$	(h,t,h)	$(0,1)$
(t,h,t)	$(1,1)$	(h,h,t)	$(0,1)$
(t,h,h)	$(0,1)$	(h,h,h)	$(0,0)$

Example: tossing 3 coins (3)

Joint pmf for X:

$$\begin{aligned} \mathbf{P}\left(X=(0,0)\right) &= \frac{1}{8}, \quad \mathbf{P}\left(X=(0,1)\right)=\frac{3}{8} \\ \mathbf{P}\left(X=(1,0)\right) &= \frac{1}{8}, \quad \mathbf{P}\left(X=(1,1)\right)=\frac{3}{8} \end{aligned}$$

Marginal pmf for X_1:

$$\begin{aligned} \mathbf{P}(X_1=0) &= \sum_{i=0}^{1} \mathbf{P}\left(X=(0,i)\right) \\ &= \mathbf{P}\left(X=(0,0)\right)+\mathbf{P}\left(X=(0,1)\right) \\ &= \frac{1}{8}+\frac{3}{8}=\frac{1}{2} \\ \mathbf{P}(X_1=1) &= \frac{1}{2} \end{aligned}$$

Example: tossing 3 coins (4)

Marginal pmf for X_2:

$$\mathbf{P}(X_2 = 0) = \frac{1}{4}, \quad \mathbf{P}(X_2 = 1) = \frac{3}{4}$$

Remark:
We have $X_1 \sim \mathcal{B}(1/2)$ and $X_2 \sim \mathcal{B}(3/4)$

Summary in a table:

$X_1 \backslash X_2$	0	1	Marg. X_1
0	1/8	3/8	1/2
1	1/8	3/8	1/2
Marg. X_2	1/4	3/4	1

Continuous case: joint density

Definition 5.

Consider the following situation:

- X, Y continuous real valued random variables

The random vector (X, Y) is said to be jointly continuous iff for "all" subsets $C \subset \mathbb{R}^2$ we have

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

Values of interest in terms of the density

Proposition 6.

Let

- X, Y random variables
- f the joint density of X, Y

Then the marginals densities of X and Y are given by

$$f_X(x) = \int_{\mathbb{R}} f(x,y)\,dy \qquad f_Y(y) = \int_{\mathbb{R}} f(x,y)\,dx$$

If $a_1 < a_2$ and $b_1 < b_2$, we also have

$$\mathbf{P}\left(a_1 < X \le a_2,\ b_1 < Y \le b_2\right) = \int_{a_1}^{a_2}\int_{b_1}^{b_2} f(x,y)\,dxdy$$

Simple example of bivariate density (1)

Density: Let (X, Y) be a random vector with density

$$2e^{-x}e^{-2y}\,\mathbf{1}_{(0,\infty)}(x)\,\mathbf{1}_{(0,\infty)}(y)$$

Question: Compute

$$\mathbf{P}(X < Y)$$

Simple example of bivariate density (2)

Computation: We have

$$\begin{aligned}
\mathbf{P}(X < Y) &= 2\int_{0<x<y<\infty} e^{-x} e^{-2y}\, dxdy \\
&= 2\int_0^\infty dy\, e^{-2y} \int_0^y e^{-x}\, dx \\
&= 2\int_0^\infty e^{-2y}(1 - e^{-y})\, dy \\
&= \frac{1}{3}
\end{aligned}$$

Change of variable in the plane (1)

Density: Let (X, Y) be a random vector with density

$$e^{-(x+y)} \mathbf{1}_{(0,\infty)}(x) \mathbf{1}_{(0,\infty)}(y)$$

Question:
Compute the density of the r.v $Z = \frac{X}{Y}$

Change of variable in the plane (2)

Characterization through expectations: Let $\varphi \in \mathcal{C}_b(\mathbb{R})$. Then

$$\mathbf{E}\left[\varphi(Z)\right] = \int_0^\infty \int_0^\infty \varphi\left(\frac{x}{y}\right) e^{-(x+y)}\, dxdy$$

Change of variable: Set

$$z = \frac{x}{y}, \quad w = y \quad \Longleftrightarrow \quad x = z\, w, \quad y = w$$

Jacobian:

$$J = w$$

Change of variable in the plane (3)

Computing $\mathbf{E}[\varphi(Z)]$:

$$\begin{aligned}\mathbf{E}[\varphi(Z)] &= \int_0^\infty \int_0^\infty \varphi(z)\, we^{-w(z+1)}\, dwdz \\ &= \int_0^\infty dz\, \varphi(z) \int_0^\infty w\, e^{-w(z+1)}\, dw \\ &= \int_0^\infty \varphi(z)\, \frac{1}{(1+z)^2}\, dz\end{aligned}$$

Density of Z:

$$\frac{1}{(1+z)^2}\, \mathbf{1}_{(0,\infty)}(z)$$

Joint cdf in higher dimensions

Definition 7.

Let

- $X_1, \ldots, X_n$ random variables
- $a_1, \ldots, a_n \in \mathbb{R}$

The following joint cdf describes
the joint distribution of $(X_1, \ldots, X_n)$:

$$F(a_1, \ldots, a_n) = \mathbf{P}\left(X_1 \leq a_1, \ldots, X_n \leq a_n\right)$$

Joint density in higher dimensions

Definition 8.

Consider the following situation:

- $X_1, \ldots, X_n$ real valued random variables

The random vector $(X_1, \ldots, X_n)$ is said to be jointly continuous iff for "all" subsets $C \subset \mathbb{R}^n$ we have

$$\mathbf{P}\left((X_1, \ldots, X_n) \in C\right) = \int_{(x_1, \ldots, x_n) \in C} f(x_1, \ldots, x_n)\, dx_1 \cdots dx_n$$

Outline

Definition of independence

Definition 9.

Let

- X, Y random variables

X and Y are said to be independent if for "all" $C, D \subset \mathbb{R}$ we have

$$\mathbf{P}(X \in C, \; Y \in D) = \mathbf{P}(X \in C)\, \mathbf{P}(Y \in D)$$

Characterizations of independence

Proposition 10.

Let X, Y random variables.
Then X and Y are independent in the following cases

1. If the joint cdf F satisfies

$$F(a, b) = F_X(a)\, F_Y(b), \quad \text{for all } a, b \in \mathbb{R}$$

2. If X, Y are discrete and the joint pmf satisfies

$$p(x, y) = p_X(x)\, p_Y(y), \quad \text{for all } (x, y) \in E_1 \times E_2$$

3. If X, Y are jointly cont. and the joint density satisfies

$$f(x, y) = f_X(x)\, f_Y(y), \quad \text{for all } (x, y) \in \mathbb{R}^2$$

Example ctd: tossing 3 coins (1)

Experiment:
Tossing a coin 3 times

Events: We consider

$$A = \text{"At most one Head"}$$
$$B = \text{"At least one Head and one Tail"}$$

Random variables: Set

$$X_1 = \mathbf{1}_A, \qquad X_2 = \mathbf{1}_B, \qquad X = (X_1, X_2)$$

Example ctd: tossing 3 coins (2)

We have seen:

$X_1 \backslash X_2$	0	1	Marg. X_1
0	1/8	3/8	1/2
1	1/8	3/8	1/2
Marg. X_2	1/4	3/4	1

Checking independence: With the help of the table, one can see that

$$\mathbf{P}(X = (i,j)) = \mathbf{P}(X_1 = i)\, \mathbf{P}(X_2 = j), \quad \text{for all} \quad i,j \in \{0,1\}$$

Therefore $X_1 \perp\!\!\!\perp X_2$.

Remark: The relation $X_1 \perp\!\!\!\perp X_2$ is due to the fact that $A \perp\!\!\!\perp B$.
$\hookrightarrow$ cf. Conditional probability, Section 4.

Example: Romeo and Juliet (1)

Situation:

- Romeo and Juliet decide to meet on the main square of Verona
- They arrive at independent times between 12pm and 1pm
- Rule: the first to arrive leaves after 10mn

Question:

Compute the probability that Romeo meets Juliet

Example: Romeo and Juliet (2)

Model:

- X = Arrival time for Romeo
- Y = Arrival time for Juliet
- Renormalize everything on $[0,1]$
- Hypothesis: $X \perp\!\!\!\perp Y$ and $X, Y \sim \mathcal{U}([0,1])$

Joint density: The joint density for (X, Y) is

$$f(x,y) = \mathbf{1}_{[0,1]^2}(x,y) = \mathbf{1}_{[0,1]}(x)\,\mathbf{1}_{[0,1]}(y)$$

Example: Romeo and Juliet (3)

Aim: Compute

$$\mathbf{P}\left(|Y - X| < \frac{1}{6}\right)$$

Complementary: Geometrically one can see that

$$\mathbf{P}\left(|Y - X| \geq \frac{1}{6}\right) = \left(\frac{5}{6}\right)^2$$

Conclusion:

$$\mathbf{P}\left(|Y - X| < \frac{1}{6}\right) = 1 - \left(\frac{5}{6}\right)^2 \simeq 30.5\%$$

Characterizations of independence

Proposition 11.

Let X, Y random variables.
Then X and Y are independent in the following cases

1. If X, Y are discrete and there exist h, g such that
$$p(x, y) = h(x)\, g(y), \quad \text{for all } (x, y) \in E_1 \times E_2$$
2. If X, Y are jointly cont. and there exist h, g such that
$$f(x, y) = h(x)\, g(y), \quad \text{for all } (x, y) \in \mathbb{R}^2$$

Example of independence (1)

Example 1: If (X, Y) have joint density

$$6e^{-(2x+3y)}\, \mathbf{1}_{(0,\infty)^2}(x, y),$$

then $X \perp\!\!\!\perp Y$.

Example of independence (2)

Recall joint density:

$$6e^{-(2x+3y)}\, \mathbf{1}_{(0,\infty)^2}(x,y)$$

Decomposition of the density:

$$f(x,y) = h(x)\, g(y),$$

with

$$h(x) = 6e^{-2x}\, \mathbf{1}_{(0,\infty)}(x), \qquad g(y) = e^{-3y}\, \mathbf{1}_{(0,\infty)}(y)$$

Conclusion:

$$X \perp\!\!\!\perp Y$$

Example of non independence (1)

Example 2: If (X, Y) have joint density

$$24xy\,\mathbf{1}_{(0,\infty)^2}(x,y)\mathbf{1}_{(0<x+y<1)},$$

then X, Y are not independent

Example of non independence (2)

Recall density:

$$f(x,y) = 24xy\, \mathbf{1}_{(0,\infty)^2}(x,y)\mathbf{1}_{(0<x+y<1)},$$

Non product structure:

X, Y satisfy the relation: $X + Y < 1$.

Checking non independence: We have

$$\mathbf{P}\left((X,Y) \in \left[0, \frac{1}{2}\right]^2\right) = \int_{[0,\frac{1}{2}]^2} 24xy\, dxdy = \frac{3}{8}$$

and

$$\mathbf{P}\left(X \in \left[0, \frac{1}{2}\right]\right)\mathbf{P}\left(Y \in \left[0, \frac{1}{2}\right]\right) = \left(24 \int_0^{\frac{1}{2}} dx\, x \int_0^{1-x} y\, dy\right)^2 = \left(\frac{11}{16}\right)^2$$

Outline

Density of a sum

Proposition 12.

Let

- X, Y continuous random variables
- Hypothesis: $X \perp\!\!\!\perp Y$
- Set $Z = X + Y$

Then the density of Z is given by

$$f_Z(a) = [f_X * f_Y](a) = \int_{\mathbb{R}} f_X(a - y) f_Y(y) \, dy$$

Proof

Characterization by expectations: Let $\varphi \in \mathcal{C}(\mathbb{R})$. Then

$$\mathbf{E}\left[\varphi(Z)\right] = \int_{\mathbb{R}^2} \varphi(x+y) f_X(x) f_Y(y)\, dxdy$$

Change of variable:
$x + y = a$ and $y = b$, thus $J = 1$

Expression for $\mathbf{E}[\varphi(Z)]$:

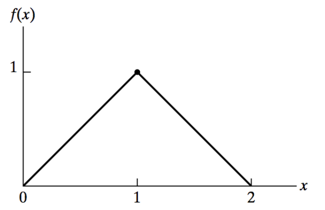
$$\mathbf{E}\left[\varphi(Z)\right] = \int_{\mathbb{R}} \varphi(a) \left(\int_{\mathbb{R}} f_X(a-b)\, f_Y(b)\, db \right) da$$

Triangular distribution

Proposition 13.

Let

- $X, Y \sim \mathcal{U}([0,1])$
- Hypothesis: $X \perp\!\!\!\perp Y$
- Set $Z = X + Y$



Then the density of Z is given by

$$f_Z(a) = a\,\mathbf{1}_{[0,1]}(a) + (2-a)\,\mathbf{1}_{[1,2]}(a)$$

Proof

Application of Proposition 12:

$$f_Z(a) = \int_0^1 f_X(a-y)\, dy = \int_{[0,1]\cap[a-1,a]} dy = |[0,1]\cap[a-1,a]|$$

Case 1: $a \in [0,1]$: Then $[0,1]\cap[a-1,a] = [0,a]$ and

$$f_Z(a) = a$$

Case 2: $a \in (1,2]$: Then $[0,1]\cap[a-1,a] = [a-1,1]$ and

$$f_Z(a) = 2-a$$

Sums of Gamma random variables

Proposition 14.

Let

- $X_1, \ldots, X_n$ independent random variables
- $X_i \sim \Gamma(t_i, \lambda)$
- $Z = \sum_{i=1}^n X_i$

Then

$$Z \sim \Gamma\left(\sum_{i=1}^n t_i, \lambda\right)$$

Remark: This result includes

- Sums of exponential random variables
- Sums of chi-square random variables

Sums of Gaussian random variables

Proposition 15.

Let

- $X_1, \ldots, X_n$ independent random variables
- $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$
- $Z = \sum_{i=1}^n X_i$

Then

$$Z \sim \mathcal{N}\left(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2\right)$$

Example: basketball (1)

Situation:

- A basketball team will play a 44-game season
- 26 games are against class A teams, with probability of win $= .4$
- 18 games are against class B teams, with probability of win $= .7$
- Results of the different games are independent.

Question: Approximate the probability that

1. The team wins 25 games or more
2. The team wins more games against class A teams than it does against class B teams

Example: basketball (2)

Model: We set

- $X_A = \#$ games the team wins against class A
- $X_B = \#$ games the team wins against class B

Then $X_A \perp\!\!\!\perp X_B$ and

$$X_A \sim \text{Bin}(26, 0.4), \quad X_B \sim \text{Bin}(18, 0.7)$$

Approximation for X_A, X_B: According to DeMoivre-Laplace,

$$X_A \approx \mathcal{N}(10.24; 6.24), \qquad X_B \approx \mathcal{N}(12.60; 3.78)$$

Example: basketball (3)

Approximation for $X_A + X_B$: Since $X_A \perp\!\!\!\perp X_B$,

$$X_A + X_B \approx \mathcal{N}(23; 10.2)$$

Question 1: We have

$$\begin{aligned} \mathbf{P}(X_A + X_B \geq 25) &= \mathbf{P}(X_A + X_B \geq 24.5) \\ &= \mathbf{P}\left(\frac{X_A + X_B - 23}{\sqrt{10.2}} \geq \frac{24.5 - 23}{\sqrt{10.2}}\right) \\ &\simeq 1 - \mathbf{P}(Z < .4739) \\ &\simeq .3178 \end{aligned}$$

Example: basketball (4)

Approximation for $X_A - X_B$: Since $X_A \perp\!\!\!\perp X_B$,

$$X_A - X_B \approx \mathcal{N}(-2.2; 10.2)$$

Question 2: We have

$$\begin{aligned}
\mathbf{P}\left(X_A - X_B > 0\right) &= \mathbf{P}\left(X_A - X_B \geq .5\right) \\
&= \mathbf{P}\left(\frac{X_A - X_B + 2.2}{\sqrt{10.2}} \geq \frac{.5 + 2.2}{\sqrt{10.2}}\right) \\
&\simeq 1 - \mathbf{P}\left(Z < .8530\right) \\
&\simeq .1968
\end{aligned}$$

Sums of Poisson random variables

Proposition 16.

Let

- $X_1, \ldots, X_n$ independent random variables
- $X_i \sim \mathcal{P}(\lambda_i)$
- $Z = \sum_{i=1}^n X_i$

Then

$$Z \sim \mathcal{P}\left(\sum_{i=1}^n \lambda_i\right)$$

Proof for 2 random variables

Hypothesis:
$X_1 \sim \mathcal{P}(\lambda_1)$, $X_2 \sim \mathcal{P}(\lambda_2)$ and $X_1 \perp\!\!\!\perp X_2$

Computation: For $n \geq 0$,

$$\begin{aligned}
\mathbf{P}\left(X_1 + X_2 = n\right) &= \sum_{k=0}^{n} \mathbf{P}\left(X_1 = k\right) \mathbf{P}\left(X_2 = n - k\right) \\
&= \sum_{k=0}^{n} e^{-\lambda_1} \frac{\lambda_1^k}{k!} \, e^{-\lambda_2} \frac{\lambda_2^{n-k}}{(n-k)!} \\
&= e^{-(\lambda_1+\lambda_2)} \frac{(\lambda_1 + \lambda_2)^n}{n!}
\end{aligned}$$

Sums of Binomial random variables

Proposition 17.

Let

- $X_1, \ldots, X_n$ independent random variables
- $X_i \sim \text{Bin}(n_i, p)$
- $Z = \sum_{i=1}^n X_i$

Then

$$Z \sim \text{Bin}\left(\sum_{i=1}^n n_i, p\right)$$

Outline

General definition

Definition 18.

Let

- (X, Y) couple of discrete random variables
- Joint pmf p
- Marginal pmf's p_X, p_Y
- y such that $p_Y(y) > 0$

Then the conditional pmf of X given $Y = y$ is defined by

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

Example ctd: tossing 3 coins (1)

Experiment:
Tossing a coin 3 times

Events: We consider

$$A = \text{"At most one Head"}$$
$$B = \text{"At least one Head and one Tail"}$$

Random variables: Set

$$X_1 = \mathbf{1}_A, \qquad X_2 = \mathbf{1}_B, \qquad X = (X_1, X_2)$$

Example ctd: tossing 3 coins (2)

We have seen:

$X_1 \backslash X_2$	0	1	Marg. X_1
0	1/8	3/8	1/2
1	1/8	3/8	1/2
Marg. X_2	1/4	3/4	1

Conditional probabilities given $X_1 = 0$:

$$p_{X_2|X_1}(0|\,0) = \frac{1/8}{1/2} = \frac{1}{4}, \quad p_{X_2|X_1}(1|\,0) = \frac{3/8}{1/2} = \frac{3}{4}$$

Conditional probabilities given $X_2 = 1$:

$$p_{X_1|X_2}(0|\,1) = \frac{3/8}{3/4} = \frac{1}{2}, \quad p_{X_1|X_2}(1|\,1) = \frac{3/8}{3/4} = \frac{1}{2}$$

Conditioning Poisson random variables

Proposition 19.

Let

- $X \sim \mathcal{P}(\lambda_1)$, $Y \sim \mathcal{P}(\lambda_2)$
- $X \perp\!\!\!\perp Y$
- $p = \frac{\lambda_1}{\lambda_1+\lambda_2}$

Then

$$\mathcal{L}\left(X | X + Y = n\right) = \text{Bin}(n, p)$$

Proof (1)

Expression for the conditional probabilities:
Let $0 \leq k \leq n$. Then invoking $X \perp\!\!\!\perp Y$,

$$\mathbf{P}\left(X = k \middle| X + Y = n\right) = \frac{\mathbf{P}\left(X = k\right)\mathbf{P}\left(Y = n - k\right)}{\mathbf{P}\left(X + Y = n\right)}$$

Law of $X + Y$: We have seen

$$X + Y \sim \mathcal{P}(\lambda_1 + \lambda_2)$$

Proof (2)

Computation of the conditional probabilities:

$$
\begin{aligned}
&\mathbf{P}\left(X = k \middle| X + Y = n\right) \\
&= e^{-\lambda_1} \frac{\lambda_1^k}{k!} e^{-\lambda_2} \frac{\lambda_2^{n-k}}{(n-k)!} \left[e^{-(\lambda_1+\lambda_2)} \frac{(\lambda_1+\lambda_2)^n}{n!} \right]^{-1} \\
&= \binom{n}{k} p^k (1-p)^{n-k}
\end{aligned}
$$

Conclusion:

$$
\mathcal{L}\left(X \middle| X + Y = n\right) = \text{Bin}(n, p)
$$

Outline

General definition

Definition 20.

Let

- (X, Y) couple of continuous random variables
- Joint density f
- Marginal densities f_X, f_Y
- y such that $f_Y(y) > 0$

Then the conditional density of X given $Y = y$ is defined by

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

Justification of the definition

Heuristics: $f_{X|Y}(x|\,y)$ can be interpreted as

$$\begin{aligned} f_{X|Y}(x|\,y)\,dx &= \frac{f(x,y)\,dxdy}{f_Y(y)\,dy} \\ &\simeq \frac{\mathbf{P}\left(x \le X \le x+dx,\; y \le Y \le y+dy\right)}{\mathbf{P}\left(y \le Y \le y+dy\right)} \\ &= \mathbf{P}\left(x \le X \le x+dx|\, y \le Y \le y+dy\right) \end{aligned}$$

Use of the conditional probability: compute probabilities like

$$\mathbf{P}\left(X \in A|\, Y = y\right) = \int_A f_{X|Y}(x|\,y)\,dx$$

Rigorous definition: see MA 539

Simple example of continuous conditioning (1)

Density: Let (X, Y) be a random vector with density

$$\frac{e^{-\frac{x}{y}} e^{-y}}{y} \mathbf{1}_{(0,\infty)}(x) \mathbf{1}_{(0,\infty)}(y)$$

Question: Compute

$$\mathbf{P}(X > 1 | Y = y)$$

Simple example of continuous conditioning (2)

Marginal distribution of Y: We have

$$
\begin{aligned}
f_Y(y) &= \int_0^\infty f(x,y)\,dx \\
&= \frac{e^{-y}}{y}\left(\int_0^\infty e^{-\frac{x}{y}}\,dx\right)\mathbf{1}_{(0,\infty)}(y) \\
&= e^{-y}\,\mathbf{1}_{(0,\infty)}(y)
\end{aligned}
$$

Conditional density: For $y > 0$ we have

$$
f_{X|Y}(x|\,y) = \frac{f(x,y)}{f_Y(y)} = \frac{e^{-\frac{x}{y}}}{y}\,\mathbf{1}_{(0,\infty)}(x)
$$

Namely $\mathcal{L}(X|\,Y = y) = \mathcal{E}(\frac{1}{y})$

Simple example of continuous conditioning (3)

Conditional probability:

$$\begin{aligned} \mathbf{P}(X > 1 | Y = y) &= \int_1^\infty f_{X|Y}(x|y)\,dx \\ &= \int_1^\infty \frac{e^{-\frac{x}{y}}}{y}\,dx \\ &= e^{-\frac{1}{y}} \end{aligned}$$

Outline

Characterizing r.v by expected values

Notation:

$C_b(\mathbb{R}^2) \equiv$ set of continuous and bounded functions on $\mathbb{R}^2$.

Theorem 21.

Let $X = (X_1, X_2)$ be a r.v in $\mathbb{R}^2$. We assume that

$$\mathbf{E}[\varphi(X_1, X_2)] = \int_{\mathbb{R}^2} \varphi(x_1, x_2) \, f(x_1, x_2) \, dx_1 \, dx_2,$$

for all functions $\varphi \in C_b(\mathbb{R}^2)$.

Then (X_1, X_2) is continuous, with density f.

Application: change of variable

Problem: Let

- $X = (X_1, X_2)$ random variable with density f.
- Set $Y = h(X)$ with $h : \mathbb{R}^2 \to \mathbb{R}^2$.

We wish to find the density of Y.

Application: change of variable (2)

Recipe: One proceeds as follows

1. For $\varphi \in C_b(\mathbb{R}^2)$, write

$$\mathbf{E}[\varphi(Y)] = \mathbf{E}[\varphi(h(X))] = \int_{\mathbb{R}^2} \varphi(h(x_1, x_2))\, f(x_1, x_2)\, dx_1 dx_2 .$$

2. Change variables $y = h(x)$ in the integral.
   After some elementary computations we get

$$\mathbf{E}[\varphi(Y)] = \int_{\mathbb{R}^2} \varphi(y_1, y_2)\, g(y_1, y_2)\, dy_1 dy_2 .$$

3. This characterizes Y, which admits a density g

Polar coordinates of Gaussian vectors (1)

Standard Gaussian vector in $\mathbb{R}^2$: Consider

- $X, Y \sim \mathcal{N}(0,1)$, with $X \perp\!\!\!\perp Y$
- $Z = (X, Y)$

Polar coordinates: Set

$$(X, Y) = (R\cos(\Theta), R\sin(\Theta))$$

Question:
Find the joint density of (R, Θ)

Polar coordinates of Gaussian vectors (2)

Decomposition of the expected value: For $\varphi \in \mathcal{C}_b(\mathbb{R}^2)$,

$$\begin{aligned} \mathbf{E}[\varphi(R,\Theta)] &= \mathbf{E}\left[\varphi(R,\Theta)\,\mathbf{1}_{(Y>0)}\right] + \mathbf{E}\left[\varphi(R,\Theta)\,\mathbf{1}_{(Y<0)}\right] \\ &\equiv A_+ + A_- \end{aligned}$$

Expression for A_+:

$$\begin{aligned} A_+ &= \mathbf{E}\left[\varphi\left((X^2+Y^2)^{1/2}, \tan^{-1}\left(\frac{Y}{X}\right)\right)\mathbf{1}_{(X>0)}\right] \\ &= \int_{\mathbb{R}\times\mathbb{R}_+} \varphi\left((x^2+y^2)^{1/2}, \tan^{-1}\left(\frac{y}{x}\right)\right) \frac{e^{-\frac{x^2+y^2}{2}}}{2\pi}\, dxdy \end{aligned}$$

Polar coordinates of Gaussian vectors (3)

Change of variable for A_+: Set

$$x = r\,\cos(\theta), \qquad y = r\,\sin(\theta) \quad \Longrightarrow \quad J(r,\theta) = r$$

Then

$$A_+ = \int_{\mathbb{R}_+\times(0,\pi)} \varphi(r,\theta)\,\frac{r\,e^{-\frac{r^2}{2}}}{2\pi}\,drd\theta$$

Change of variable for A_-: We find

$$A_- = \int_{\mathbb{R}_+\times(\pi,2\pi)} \varphi(r,\theta)\,\frac{r\,e^{-\frac{r^2}{2}}}{2\pi}\,drd\theta$$

Polar coordinates of Gaussian vectors (4)

Expression for the expected value:

$$\mathbf{E}\left[\varphi(R,\Theta)\right] = \int_{\mathbb{R}_+\times(0,2\pi)} \varphi(r,\theta)\, \frac{r\, e^{-\frac{r^2}{2}}}{2\pi}\, dr d\theta$$

Joint density for (R,Θ):

$$f(r,\theta) = \frac{1}{2\pi}\,\mathbf{1}_{(0,2\pi)}(\theta) \times r\, e^{-\frac{r^2}{2}}\,\mathbf{1}_{\mathbb{R}_+}(r)$$

Otherwise stated:

- $R \sim$ Rayleigh, $\Theta \sim \mathcal{U}([0,2\pi])$
- $R \perp\!\!\!\perp \Theta$

Change of variable: general result

Theorem 22.

Let

- $X = (X_1, X_2)$ continuous random variable
- Density: f_X
- g diffeomorphism of $\mathbb{R}^2$
- $Y = g(X)$

Then Y has a density f_Y given by

$$f_Y(y) = f_X\left(g^{-1}(y)\right) J(y)\, \mathbf{1}_{\{y=g(x) \text{ for some } x\}}$$

Outline

Cond. pmf in the discrete case (repeated)

Definition 23.

Let

- (X, Y) couple of discrete random variables
- Joint pmf p
- Marginal pmf's p_X, p_Y
- y such that $p_Y(y) > 0$

Then the conditional pmf of X given $Y = y$ is defined by

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

Cond. expectation in the discrete case

Definition 24.

Let

- (X, Y) couple of discrete random variables
- Joint pmf p
- Marginal pmf's p_X, p_Y, y such that $p_Y(y) > 0$
- $p_{X|Y}(x|\, y)$ conditional distribution

Then the conditional exp. of X given $Y = y$ is defined by

$$\mathbf{E}\left[X|\, Y = y\right] = \sum_{x \in \mathcal{E}} x\, p_{X|Y}(x|\, y)$$

Binomial example (1)

Situation: Let

- $X, Y \sim \text{Bin}(n, p)$
- $Z = X + Y$

Problem: We wish to compute

$$\mathbf{E}[X | Z = m]$$

Binomial example (2)

Distribution for Z:

$$Z = \sum_{i=1}^{n} X_i + \sum_{j=1}^{n} Y_j \sim \text{Bin}(2n, p)$$

Computation for conditional pmf: For $k \leq \min(n, m)$ we have

$$\begin{aligned}
\mathbf{P}\left(X = k \middle| Z = m\right) &= \frac{\mathbf{P}(X = k,\, X + Y = m)}{\mathbf{P}(Z = m)} \\
&= \frac{\mathbf{P}(X = k,\, Y = m - k)}{\mathbf{P}(Z = m)} \\
&= \frac{\binom{n}{k}\binom{n}{m-k}}{\binom{2n}{m}}
\end{aligned}$$

Binomial example (3)

Conditional pmf: For $k \leq \min(n, m)$ we have

$$p_{X|Z}(k|\, m) = \frac{\binom{n}{k}\binom{n}{m-k}}{\binom{2n}{m}}$$

Recall: If $V \sim \text{HypG}(n, N, m)$ then

$$\mathbf{P}(X = k) = \frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}}$$

Identification of the conditional pmf: We have

$$p_{X|Z}(k|\, m) = \text{Pmf of HypG}(2n, m, n)$$

Binomial example (4)

Conditional expectation: Let $V \sim \text{HypG}(2n, m, n)$. Then

$$\mathbf{E}[X|\, Z = m] = \mathbf{E}[V]$$

Numerical value:
According to the values for hypergeometric distributions,

$$\mathbf{E}[X|\, Z = m] = m \times \frac{n}{2n} = \frac{m}{2}$$

Cond. density in the continuous case (repeated)

Definition 25.

Let

- (X, Y) couple of continuous random variables
- Joint density f
- Marginal densities f_X, f_Y
- y such that $f_Y(y) > 0$

Then the conditional density of X given $Y = y$ is defined by

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

Cond. expectation in the continuous case

Definition 26.

Let

- (X, Y) couple of continuous random variables
- Joint density f
- Marginal densities f_X, f_Y, y such that $f_Y(y) > 0$
- $f_{X|Y}(x|\, y)$ conditional density

Then the conditional exp. of X given $Y = y$ is defined by

$$\mathbf{E}\left[X|\, Y = y\right] = \int_{\mathbb{R}} x\, f_{X|Y}(x|\, y)\, dx$$

Example of continuous conditional expectation (1)

Density: Let (X, Y) be a random vector with density

$$\frac{e^{-\frac{x}{y}} e^{-y}}{y} \mathbf{1}_{(0,\infty)}(x) \mathbf{1}_{(0,\infty)}(y)$$

Question: Compute

$$\mathbf{E}[X | Y = y]$$

Example of continuous conditional expectation (2)

Conditional density: For $y > 0$ we have seen that

$$f_{X|Y}(x|\,y) = \frac{f(x,y)}{f_Y(y)} = \frac{e^{-\frac{x}{y}}}{y}\,\mathbf{1}_{(0,\infty)}(x)$$

Namely $\mathcal{L}(X|\,Y=y) = \mathcal{E}(\frac{1}{y})$

Conditional expectation: We have

$$\begin{aligned} \mathbf{E}[X|\,Y=y] &= \int_{\mathbb{R}} x\, f_{X|Y}(x|\,y)\,dx \\ &= \int_0^\infty x\,\frac{e^{-\frac{x}{y}}}{y} \\ &= y \end{aligned}$$

Expectation and conditioning

Proposition 27.

Let X, Y be two random variables. Then

1. If X, Y are discrete we have
$$\mathbf{E}[X] = \sum_y \mathbf{E}\left[X \middle| Y = y\right] p_Y(y)$$
2. If X, Y are continuous we have
$$\mathbf{E}[X] = \int_{\mathbb{R}} \mathbf{E}\left[X \middle| Y = y\right] f_Y(y)\, dy$$
3. Unified notation:
$$\mathbf{E}[X] = \mathbf{E}\left\{\mathbf{E}\left[X \middle| Y\right]\right\}$$

Example: sales in a store (1)

Situation:

We consider a store on a given day. We assume

- # of people entering in the store has mean 50
- Amount of money spent by each person is $8
- Indep. between # persons entering and amount of money spent

Question:

Expected amount of money spent in the store on a given day?

Example: sales in a store (2)

Notation: We set

- $N = \#$ of customers entering the store
- $X_i =$ Amount spent by i-th customer, for $i \geq 1$
- $Z =$ Total amount spent

Expression for Z: We have (double randomness)

$$Z = \sum_{i=1}^{N} X_i$$

Hypothesis:

- X_i's follow the same distribution X
- $(X_i)_{i \geq 1} \perp\!\!\!\perp N$

Example: sales in a store (3)

Computation:

$$
\begin{aligned}
\mathbf{E}[Z] &= \mathbf{E}\left\{\mathbf{E}\left[\sum_{i=1}^{N} X_i \mid N\right]\right\} \\
&= \sum_{n=1}^{\infty} \mathbf{E}\left[\sum_{i=1}^{N} X_i \mid N = n\right] p_N(n) \\
&= \sum_{n=1}^{\infty} \mathbf{E}\left[\sum_{i=1}^{n} X_i \mid N = n\right] p_N(n) \\
&= \sum_{n=1}^{\infty} \sum_{i=1}^{n} \mathbf{E}\left[X_i \mid N = n\right] p_N(n) \\
&= \sum_{n=1}^{\infty} n\, \mathbf{E}[X]\, p_N(n) \\
&= \mathbf{E}[N]\, \mathbf{E}[X]
\end{aligned}
$$