Proability and Statistics

MA20205

Bibhas Adhikari

Autumn 2022-23, IIT Kharagpur

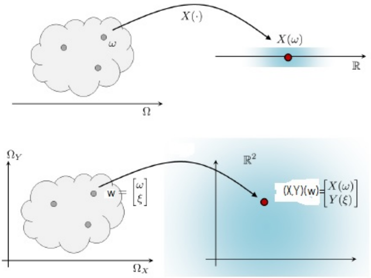
Lecture 11
October 10, 2022

Joint distributions

- A single variable X is described by a one-variable pdf $f(x)$

Joint distributions

- A single variable X is described by a one-variable pdf $f(x)$
- A pair of random variables (X, Y) is described by a two-variable pdf $f(x, y)$



Joint distributions

Let X and Y be two random variables with sample spaces Ω_X and Ω_Y respectively.
Then the joint random variable is given by $(X, Y) : \Omega_X \times \Omega_y \to \mathbb{R} \times \mathbb{R}$

Joint pmf for pair of discrete rvs

Let X and Y be two discrete random variables. The joint pmf of (X, Y) is defined as

$$f(x, y) = P(X = x \text{ and } Y = y) = P((\omega, \eta) : X(\omega) = x \text{ and } Y(\zeta) = y)$$

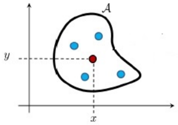
Joint distributions

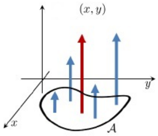
Let X and Y be two random variables with sample spaces Ω_X and Ω_Y respectively.
Then the joint random variable is given by $(X, Y) : \Omega_X \times \Omega_y \to \mathbb{R} \times \mathbb{R}$

Joint pmf for pair of discrete rvs

Let X and Y be two discrete random variables. The joint pmf of (X, Y) is defined as

$$f(x, y) = P(X = x \text{ and } Y = y) = P((\omega, \eta) : X(\omega) = x \text{ and } Y(\zeta) = y)$$





Joint distributions

Let X be a random variable for a coin toss and Y be draw of a die. The sample space is $\Omega_X \times \Omega_Y = \{(\omega, \eta) : \omega \in \{0, 1\}, \eta \in \{1, 2, 3, 4, 5, 6\}\}$. Then

$$f(x, y) = \frac{1}{12}, (x, y) \in S$$

is a pmf corresponding to (X, Y).

Joint distributions

Let X be a random variable for a coin toss and Y be draw of a die. The sample space is $\Omega_X \times \Omega_Y = \{(\omega, \eta) : \omega \in \{0, 1\}, \eta \in \{1, 2, 3, 4, 5, 6\}\}$. Then

$$f(x, y) = \frac{1}{12}, (x, y) \in S$$

is a pmf corresponding to (X, Y).

Questions

1. Let $A = \{(x, y) : x + y = 3\}$. Then $P(A) =?$

Joint distributions

Let X be a random variable for a coin toss and Y be draw of a die. The sample space is $\Omega_X \times \Omega_Y = \{(\omega, \eta) : \omega \in \{0,1\}, \eta \in \{1,2,3,4,5,6\}\}$. Then

$$f(x,y) = \frac{1}{12}, (x,y) \in S$$

is a pmf corresponding to (X, Y).

Questions

1. Let $A = \{(x,y) : x + y = 3\}$. Then $P(A) =?$
2. Let $B = \{(x,y) : \min\{x,y\} = 1\}$. Then $P(B) =?$

Joint distributions

Joint pdf for continuous rvs

Let X, Y be continuous rvs with sample spaces Ω_X, Ω_Y respectively. Then joint pdf of (X, Y) is a function $f(x, y)$ such that

$$P(A) = \int_A f(x, y)dxdy$$

for any event $A \subset \Omega_X \times \Omega_Y$.

Joint distributions

Joint pdf for continuous rvs

Let X, Y be continuous rvs with sample spaces Ω_X, Ω_Y respectively. Then joint pdf of (X, Y) is a function $f(x, y)$ such that

$$P(A) = \int_A f(x, y) dxdy$$

for any event $A \subset \Omega_X \times \Omega_Y$.

For example, if $A = [a, b] \times [c, d]$ then

$$P(A) = P(a \leq X \leq b, c \leq Y \leq d) = \int_c^d \int_a^b f(x, y) dxdy$$

Joint distributions

Joint pdf for continuous rvs

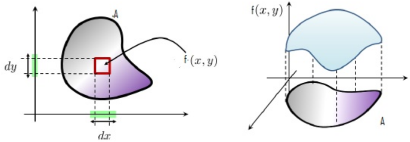
Let X, Y be continuous rvs with sample spaces Ω_X, Ω_Y respectively. Then joint pdf of (X, Y) is a function $f(x, y)$ such that

$$P(A) = \int_A f(x, y)dxdy$$

for any event $A \subset \Omega_X \times \Omega_Y$.

For example, if $A = [a, b] \times [c, d]$ then

$$P(A) = P(a \leq X \leq b, c \leq Y \leq d) = \int_c^d \int_a^b f(x, y)dxdy$$



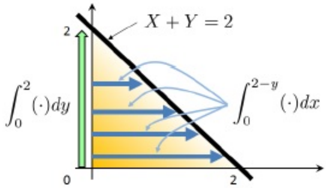
Joint distributions

Let (X, Y) be a joint random variable with uniform distribution on $[0, 2] \times [0, 2]$. Then find $P(A)$ if $A = \{(x, y) : x + y \leq 2\}$.

Joint distributions

Let (X, Y) be a joint random variable with uniform distribution on $[0,2] \times [0,2]$. Then find $P(A)$ if $A = \{(x,y) : x + y \leq 2\}$.

$$\begin{aligned} P(A) &= \int_A f(x,y)\, dx\, dy \\ &= \int_0^2 \int_0^{2-y} \frac{1}{4} dxdy \\ &= \frac{1}{2}. \end{aligned}$$



Joint distributions

Normalization property Let $\Omega = \Omega_X \times \Omega_Y$. All joint pmfs and pdfs satisfy

$$\sum_{(x,y)\in\Omega} f(x,y) = 1 \text{ or } \int_{\Omega} f(x,y)dxdy = 1.$$

Joint distributions

Normalization property Let $\Omega = \Omega_X \times \Omega_Y$. All joint pmfs and pdfs satisfy

$$\sum_{(x,y)\in\Omega} f(x,y) = 1 \text{ or } \int_\Omega f(x,y)dxdy = 1.$$

Problem Find the value of k for which

$$f(x,y) = \begin{cases} ke^{-x}e^{-y} \; 0 \le y \le x < \infty \\ 0 \text{ otherwise} \end{cases}$$

is a joint pdf.

Joint distributions

Marginal pmf and pdf

The marginal pmf is defined by

$$f_X(x) = \sum_{y\in\Omega_Y} f(x,y),\ f_Y(y) = \sum_{x\in\Omega_X} f(x,y)$$

and the marginal pdf is defined as

$$f_X(x) = \int_{y\in\Omega_Y} f(x,y)dy,\ f_Y(y) = \int_{x\in\Omega_X} f(x,y)dx$$

Joint distributions

Marginal pmf and pdf

The marginal pmf is defined by

$$f_X(x) = \sum_{y\in\Omega_Y} f(x,y),\ f_Y(y) = \sum_{x\in\Omega_X} f(x,y)$$

and the marginal pdf is defined as

$$f_X(x) = \int_{y\in\Omega_Y} f(x,y)dy,\ f_Y(y) = \int_{x\in\Omega_X} f(x,y)dx$$

Joint Gaussian/normal random variable

A joint Gaussioan random variable (X, Y) has a joint pdf given by

$$f(x,y) = \frac{1}{2\pi\sigma^2}\exp\left\{-\frac{(x-\mu_x)^2+(y-\mu_Y)^2}{2\sigma^2}\right\}$$

Joint distributions

Marginal pdfs of Gaussian:

$$
\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f(x,y)dy \\
&= \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{(x-\mu_X)^2+(y-\mu_Y)^2}{2\sigma^2}\right\} \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu_X)^2}{2\sigma^2}\right\} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(y-\mu_Y)^2}{2\sigma^2}\right\} dy
\end{aligned}
$$

Joint distributions

Marginal pdfs of Gaussian:

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f(x,y)dy \\ &= \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{(x-\mu_X)^2+(y-\mu_Y)^2}{2\sigma^2}\right\} \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu_X)^2}{2\sigma^2}\right\} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(y-\mu_Y)^2}{2\sigma^2}\right\} dy \end{aligned}$$

Thus

$$\begin{aligned} f_X(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu_X)^2}{2\sigma^2}\right\} \\ f_Y(y) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(y-\mu_Y)^2}{2\sigma^2}\right\} \end{aligned}$$

Joint distributions

From the above derivation, note that

$$f(x, y) = f_X(x) f_Y(y).$$

Joint distributions

From the above derivation, note that

$$f(x, y) = f_X(x) f_Y(y).$$

Independent random variables

Random variables X and Y are independent if and only if

$$f(x, y) = f_X(x) f_Y(y)$$

Joint distributions

From the above derivation, note that

$$f(x, y) = f_X(x)f_Y(y).$$

Independent random variables

Random variables X and Y are independent if and only if

$$f(x, y) = f_X(x)f_Y(y)$$

QuestionHow to justify this definition?

Joint distributions

From the above derivation, note that

$$f(x, y) = f_X(x)f_Y(y).$$

Independent random variables

Random variables X and Y are independent if and only if

$$f(x, y) = f_X(x)f_Y(y)$$

QuestionHow to justify this definition?

Sequence of independent random variables

A sequence of random variables $X_1, X_2, \ldots, X_n$ is independent if and only if

$$f(x_1, x_2, \ldots, x_n) = f_{X_1}(x_1)f_{X_2}(x_2)\ldots f_{X_n}(x_n)$$

Joint distributions

From the above derivation, note that

$$f(x, y) = f_X(x)f_Y(y).$$

Independent random variables

Random variables X and Y are independent if and only if

$$f(x, y) = f_X(x)f_Y(y)$$

QuestionHow to justify this definition?

Sequence of independent random variables

A sequence of random variables $X_1, X_2, \ldots, X_n$ is independent if and only if

$$f(x_1, x_2, \ldots, x_n) = f_{X_1}(x_1)f_{X_2}(x_2)\ldots f_{X_n}(x_n)$$

Joint distributions

Terminology If n random variables are independent and have the same distribution, then the random variables are called independent and identically distributed (iid) random variables.

Joint distributions

Terminology If n random variables are independent and have the same distribution, then the random variables are called independent and identically distributed (iid) random variables.

Joint cdf

Let X and Y be random variables. The joint cdf of X and Y is defined as

$$F(x, y) = P(X \leq x \cap Y \leq y).$$

Joint distributions

Terminology If n random variables are independent and have the same distribution, then the random variables are called independent and identically distributed (iid) random variables.

Joint cdf

Let X and Y be random variables. The joint cdf of X and Y is defined as

$$F(x,y) = P(X \leq x \cap Y \leq y).$$

Obviously.

$$F(x,y) = \sum_{y' \leq y} \sum_{x' \leq x} f(x', y')$$

if X and Y are discrete, and

$$F(x,y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f(x,y)\, dx\, dy$$

Joint distributions

Observations

1. If X and Y are independent then

$$F(x, y) = F_X(x)F_Y(y)$$

Joint distributions

Observations

1. If X and Y are independent then

$$F(x, y) = F_X(x)F_Y(y)$$

2. Let X and Y are iid with pdf Unif$(0, 1)$. Then

$$F(x, y) = xy$$

Joint distributions

Observations

1. If X and Y are independent then

$$F(x, y) = F_X(x)F_Y(y)$$

2. Let X and Y are iid with pdf $\text{Unif}(0, 1)$. Then

$$F(x, y) = xy$$

3. If X and Y are iid with pdf $\mathcal{N}(\mu, \sigma^2)$ then

$$F(x, y) = \Phi\left(\frac{x-\mu}{\sigma}\right)\Phi\left(\frac{y-\mu}{\sigma}\right)$$

Joint distributions

Observations

1. $F(x, -\infty) = 0$

Joint distributions

Observations

1. $F(x, -\infty) = 0$
2. $F(-\infty, y) = 0$

Joint distributions

Observations

1. $F(x, -\infty) = 0$
2. $F(-\infty, y) = 0$
3. $F(-\infty, -\infty) = 0$

Joint distributions

Observations

1. $F(x, -\infty) = 0$
2. $F(-\infty, y) = 0$
3. $F(-\infty, -\infty) = 0$
4. $F(\infty, \infty) = 1$

Joint distributions

Observations

1. $F(x, -\infty) = 0$
2. $F(-\infty, y) = 0$
3. $F(-\infty, -\infty) = 0$
4. $F(\infty, \infty) = 1$

Marginal cdf

The marginal cdfs are

$$F_X(x) = F(x, \infty), \; F_Y(y) = F(\infty, y)$$

Joint distributions

Question How to obtain joint pdf from joint cdf?

$$f(x,y) = \frac{\partial^2}{\partial y\,\partial x} F(x,y) = \frac{\partial^2}{\partial x\,\partial y} F(x,y)$$

Joint distributions

Question How to obtain joint pdf from joint cdf?

$$f(x, y) = \frac{\partial^2}{\partial y\, \partial x} F(x, y) = \frac{\partial^2}{\partial x\, \partial y} F(x, y)$$

Example Let X and Y be random variables with joint cdf

$$F(x, y) = (1 - e^{-\lambda x})(1 - e^{-\lambda y}), x \geq 0, y \geq 0$$

Then

Joint distributions

Question How to obtain joint pdf from joint cdf?

$$f(x, y) = \frac{\partial^2}{\partial y\, \partial x} F(x, y) = \frac{\partial^2}{\partial x\, \partial y} F(x, y)$$

Example Let X and Y be random variables with joint cdf

$$F(x, y) = (1 - e^{-\lambda x})(1 - e^{-\lambda y}), x \geq 0, y \geq 0$$

Then

$$f(x, y) = \lambda^2 e^{-\lambda x} e^{-\lambda y}$$

Joint distributions

Joint expectation

Let X, Y be random variables. Then the joint expectation of the pair is defined as

$$E(XY) = \begin{cases} \sum_{y\in\Omega_Y} \sum_{x\in\Omega_X} xy\, f(x,y) \text{ if } X, Y \text{ are discrete} \\ \\ \int_{y\in\Omega_Y} \int_{x\in\Omega_X} xy\, f(x,y)\, dx\, dy \text{ if } X, Y \text{ are continuous} \end{cases}$$

Joint distributions

Joint expectation

Let X, Y be random variables. Then the joint expectation of the pair is defined as

$$E(XY) = \begin{cases} \sum_{y\in\Omega_Y} \sum_{x\in\Omega_X} xy\, f(x,y) \text{ if } X, Y \text{ are discrete} \\ \int_{y\in\Omega_Y} \int_{x\in\Omega_X} xy\, f(x,y)\, dx\, dy \text{ if } X, Y \text{ are continuous} \end{cases}$$

Question Why is the joint expectation defined as the product random variable instead of addition $(E(X+Y))$ or difference $(E(X-Y))$ or quotient $(E(X/Y))$?

Joint distributions

Suppose X and Y are discrete random variables with range spaces $\Omega_X = \{x_1, \ldots, x_n\}$ and $\Omega_Y = \{y_1, \ldots, y_n\}$.

Joint distributions

Suppose X and Y are discrete random variables with range spaces $\Omega_X = \{x_1, \ldots, x_n\}$ and $\Omega_Y = \{y_1, \ldots, y_n\}$. Now define the vectors

$$\boldsymbol{x} = [x_1, \ldots, x_n]^T, \; \boldsymbol{y} = [y_1, \ldots, y_n]^T.$$

Joint distributions

Suppose X and Y are discrete random variables with range spaces $\Omega_X = \{x_1, \ldots, x_n\}$ and $\Omega_Y = \{y_1, \ldots, y_n\}$. Now define the vectors

$$\boldsymbol{x} = [x_1, \ldots, x_n]^T, \ \boldsymbol{y} = [y_1, \ldots, y_n]^T.$$

Then define the pmf matrix as

$$\boldsymbol{P} = \begin{bmatrix} f(x_1, y_1) & f(x_1, y_2) & \ldots & f(x_1, y_n) \\ f(x_2, y_1) & f(x_2, y_2) & \ldots & f(x_2, y_n) \\ \vdots & \vdots & \ddots & \vdots \\ f(x_n, y_1) & f(x_n, y_2) & \ldots & f(x_n, y_n) \end{bmatrix}.$$

Joint distributions

Suppose X and Y are discrete random variables with range spaces $\Omega_X = \{x_1, \ldots, x_n\}$ and $\Omega_Y = \{y_1, \ldots, y_n\}$. Now define the vectors

$$\boldsymbol{x} = [x_1, \ldots, x_n]^T, \ \boldsymbol{y} = [y_1, \ldots, y_n]^T.$$

Then define the pmf matrix as

$$\boldsymbol{P} = \begin{bmatrix} f(x_1, y_1) & f(x_1, y_2) & \ldots & f(x_1, y_n) \\ f(x_2, y_1) & f(x_2, y_2) & \ldots & f(x_2, y_n) \\ \vdots & \vdots & \ddots & \vdots \\ f(x_n, y_1) & f(x_n, y_2) & \ldots & f(x_n, y_n) \end{bmatrix}.$$

Then

$$E(XY) = \boldsymbol{x}^T \boldsymbol{P} \boldsymbol{y},$$

the weighted inner (scalar) product of $\boldsymbol{x}$ and $\boldsymbol{y}$.

Joint distributions

For example, if $\Omega_X = \{1, \ldots, n\} = \Omega_Y$ with

$$f(x,y) = \begin{cases} \frac{1}{n} x = y \\ 0\, x \neq y \end{cases}$$

then

$$\boldsymbol{P} = \frac{1}{n}\boldsymbol{I},\ E(XY) = \frac{1}{n}\boldsymbol{x}^T\boldsymbol{y}$$

Joint distributions

Recall that the cosine angle between $\boldsymbol{x}$ and $\boldsymbol{y}$ is defined as

$$\cos\theta = \frac{\boldsymbol{x}^T\boldsymbol{y}}{\|\boldsymbol{x}\|\,\|\boldsymbol{y}\|}$$

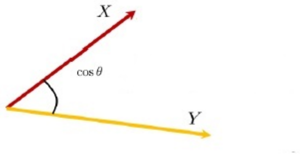
where $\|\boldsymbol{x}\| = \sqrt{\sum_{i=1}^n x_i^2}$ and $\|\boldsymbol{y}\| = \sqrt{\sum_{i=1}^n y_i^2}$.

Joint distributions

Recall that the cosine angle between $\boldsymbol{x}$ and $\boldsymbol{y}$ is defined as

$$\cos\theta = \frac{\boldsymbol{x}^T\boldsymbol{y}}{\|\boldsymbol{x}\|\,\|\boldsymbol{y}\|}$$

where $\|\boldsymbol{x}\| = \sqrt{\sum_{i=1}^n x_i^2}$ and $\|\boldsymbol{y}\| = \sqrt{\sum_{i=1}^n y_i^2}$.
Geometry of expectation: geometry defined by the weighted inner product and weighted norm



where

$$\cos\theta = \frac{\boldsymbol{x}^T\boldsymbol{P}\boldsymbol{y}}{\|\boldsymbol{x}\|_{\boldsymbol{P}_X}\,\|\boldsymbol{y}\|_{\boldsymbol{P}_Y}} = \frac{E(XY)}{\sqrt{E(X^2)}\,\sqrt{E(Y^2)}}$$

Joint distribution

In the above,

$$
\begin{aligned}
E(X^2) &= \boldsymbol{x}^T \boldsymbol{P}_X \boldsymbol{x} = \|\boldsymbol{x}\|^2_{\boldsymbol{P}_X} \\
E(Y^2) &= \boldsymbol{y}^T \boldsymbol{P}_Y \boldsymbol{y} = \|\boldsymbol{y}\|^2_{\boldsymbol{P}_Y}
\end{aligned}
$$

where

$$
\boldsymbol{P}_X = \begin{bmatrix} p(x_1) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & p(x_n) \end{bmatrix}, \boldsymbol{P}_Y = \begin{bmatrix} p(y_1) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & p(y_n) \end{bmatrix}
$$

Joint distribution

In the above,

$$\begin{aligned} E(X^2) &= \boldsymbol{x}^T \boldsymbol{P}_X \boldsymbol{x} = \|\boldsymbol{x}\|^2_{\boldsymbol{P}_X} \\ E(Y^2) &= \boldsymbol{y}^T \boldsymbol{P}_Y \boldsymbol{y} = \|\boldsymbol{y}\|^2_{\boldsymbol{P}_Y} \end{aligned}$$

where

$$\boldsymbol{P}_X = \begin{bmatrix} p(x_1) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & p(x_n) \end{bmatrix}, \boldsymbol{P}_Y = \begin{bmatrix} p(y_1) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & p(y_n) \end{bmatrix}$$

Obviously,

$$-1 \leq \frac{E(XY)}{\sqrt{E(X^2)}\sqrt{E(Y^2)}} \leq 1$$

due to Cauchy-Schwarz inequality:

$$(E(XY))^2 \leq E(X^2)E(Y^2)$$