Random Variables

CSGE602013 –STATISTICS AND PROBABILITY

FACULTY OF COMPUTER SCIENCE UNIVERSITAS INDONESIA

References

- Introduction to Probability and Statistics for Engineers & Scientists, 4th ed., Sheldon M. Ross, Elsevier, 2009.
- A Modern Introduction to Probability and Statistics, Understanding Why and How, Frederik Michel Dekking et al., Springer, 2005.

Tossing Two Die...

- E = the event that the sum of both die equals 7

  $E = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$

- We are interested in the sum of the two dice

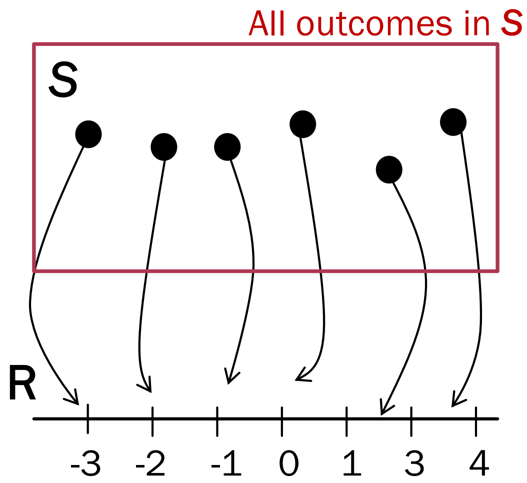
  ....We don't care about the values of the individual dice

We are often NOT interested in all of the details of the experimental result but only in the value of some numerical quantity determined by the result.

Random Variables

- Random variable X: a function that assigns a number $X(s)$ to each outcome of an experiment.
- Gives us the power of abstraction and allows us to discard unimportant details of outcomes in an experiment.

$$X : S \rightarrow R \qquad (or\ X(s) \in R, \forall s \in S)$$



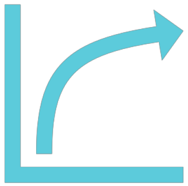
Types of Random Variables

Discrete Random Variable

Takes only finite or countably infinite number of values

Arise from discrete event measurement

Ex: sequences



Continuous Random Variable

Takes infinite and uncountable number of values

Pertinent to continuous measurement

Ex: an interval

DISCRETE RANDOM VARIABLES

Tossing two die..

1, 1	2, 1	3, 1	4, 1	5, 1	6, 1
1, 2	2, 2	3, 2	4, 2	5, 2	6, 2
1, 3	2, 3	3, 3	4, 3	5, 3	6, 3
1, 4	2, 4	3, 4	4, 4	5, 4	6, 4
1, 5	2, 5	3, 5	4, 5	5, 5	6, 5
1, 6	2, 6	3, 6	4, 6	5, 6	6, 6

Tossing two die..

- Let X be the random variable defined as the **sum of two fair dice**
- X is a discrete random variable
    - $X = 2 \rightarrow \{(1,1)\}$
    - $X = 3 \rightarrow \{(1,2),(2,1)\}$
    - $X = 4 \rightarrow \{(1,3),(2,2),(3,1)\}$
    - $X = 5 \rightarrow \{(1,4),(2,3),(3,2),(4,1)\}$
    - …
    - $X = 11 \rightarrow \{(5,6),(6,5)\}$
    - $X = 12 \rightarrow \{(6,6)\}$

Probability Mass Function (PMF)

- The probability mass function $p(x_i)$ of X is defined by

$$p(x_i) = P(X = x_i)$$

  - where $P(X = x_i)$: the probability that the value of X is equal to x_i.
    - $P(X = 2) = P(\{(1,1)\}) = 1/36$
    - $P(X = 3) = P(\{(1,2),(2,1)\}) = 2/36$
    - $P(X = 4) = P(\{(1,3),(2,2),(3,1)\}) = 3/36$
    - $P(X = 5) = P(\{(1,4),(2,3),(3,2),(4,1)\}) = 4/36$
    - …
    - $P(X = 11) = P(\{(5,6),(6,5)\}) = 2/36$
    - $P(X = 12) = P(\{(6,6)\}) = 1/36$

Conditions of PMF

1. If X must assume one of the values x_1, x_2, ..., then,

$$\begin{aligned} &0 \le p(x) \le 1 && \textit{for all real } x \\ &p(x_i) > 0, && i = 1,2,\ldots \\ &p(x) = 0, && \textit{all other values of } x \end{aligned}$$

2. Set of the values x_1, x_2, ... can be **finite** or **countably infinite.**

Conditions of PMF (2)

3. The following property must hold

$$\sum_{x_i} p(x_i) = 1$$

or, if X takes value from **finite** set {x_1, x_2, ..., x_n}, then

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

since

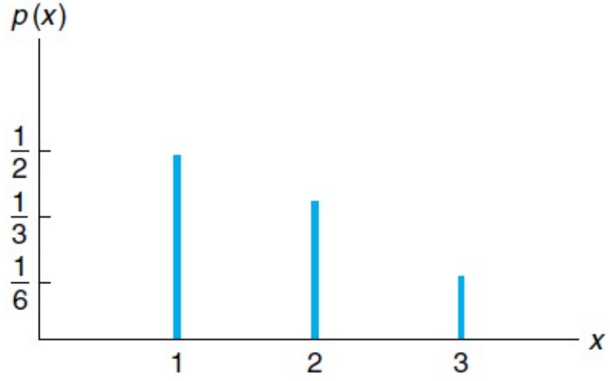
$$1 = P(S) = P(\bigcup_{i=1}^{n} X = x_i) = \sum_{i=1}^{n} P(X = x_i)$$

Contoh

- Diketahui nilai variabel acak **X** sama dengan 1, 2, atau 3.
- Misalkan, diberikan PMF dari X sebagai berikut
  - $\boldsymbol{P}(\boldsymbol{X} = 1) = p(1) = 1/2$
  - $\boldsymbol{P}(\boldsymbol{X} = 2) = p(2) = 1/3$
  - $\boldsymbol{P}(\boldsymbol{X} = 3) = p(3) = 1/6$
- PMF dari X dapat ditulis sebagai berikut:

x_i	1	2	3
$p(x_i)$	$1/2$	$1/3$	$1/6$

*) $\sum_{i=1}^{3} P(X = x_i) = \frac{1}{2} + \frac{1}{3} + \frac{1}{6} = 1$



Cumulative Distribution Function (CDF)

- Cumulative Distribution Function, F of the random variable X is defined for any real number x by

$$F_X(x) = P(X \leq x) \qquad x \in R$$

- How can we determine $P(a < X \leq b)$?

$$P(X \leq b) = P(X \leq a) + P(a < X \leq b)$$

$$\begin{aligned} P(a < X \leq b) & \\ &= P(X \leq b) - P(X \leq a) \\ &= F_X(b) - F_X(a) \end{aligned}$$

  - CDF is also called by "**distribution function**"

Basic Properties of CDF

- CDF is a monotone increasing function of X

$$F_X(t_1) \leq F_X(t_2), \qquad t_1 \leq t_2$$

- Limit $x \to \infty$ untuk F_X

$$\lim_{x \to -\infty} F_X(x) = 0 \qquad \lim_{x \to +\infty} F_X(x) = 1$$

- CDF is right-continuous

$$\lim_{\varepsilon \to 0^+} F_X(x + \varepsilon) = F_X(x)$$

CDF for Discrete RV

- By definition of CDF:

$$F(t) = P(X \le t)$$

- In discrete case, it becomes

$$\begin{aligned} F(t) &= P(X \le t) \\ &= \sum_{x_i \le t} P(X = x_i) \\ &= \sum_{x_i \le t} p(x_i) \end{aligned}$$

  - we can omit subscript $\boldsymbol{X}$ on $\boldsymbol{F_X}(t)$ since it's clear here that $\boldsymbol{F}(t)$ is CDF of RV $\boldsymbol{X}$

Creating CDF of a Discrete RV

- Suppose we have discrete R.V. $\boldsymbol{X}$ and its Probability Mass Function (PMF) as follows:

x_i	1	2	3
$p(x_i)$	$1/2$	$1/3$	$1/6$

- Determine the Cumulative Distribution Function (CDF) of random variable $\boldsymbol{X}$!

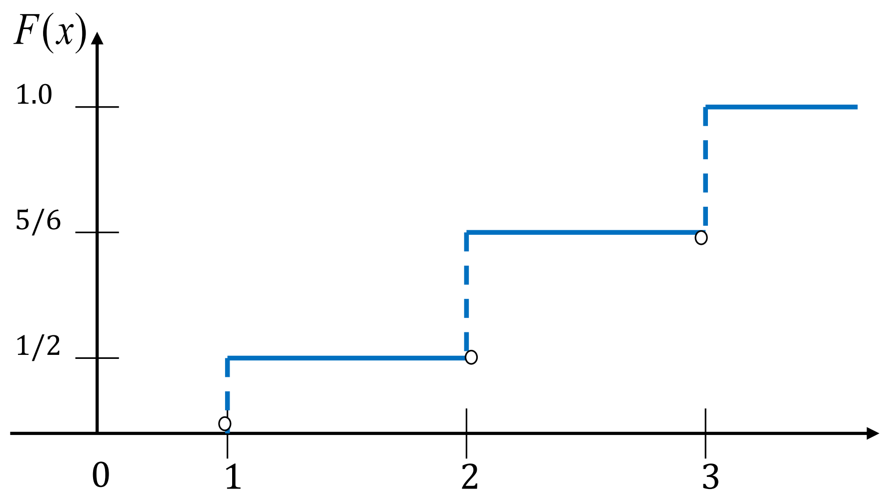
$$F_X(a) = \begin{cases} 0 & a < 1 \\ 1/2 & 1 \le a < 2 \\ 1/2 + 1/3 = 5/6 & 2 \le a < 3 \\ 1/2 + 1/3 + 1/6 = 1 & a \ge 3 \end{cases}$$

Although **X** is discrete, CDF must be defined on **real values.**

Creating CDF of Discrete RV (2)

- Draw the **graph** of CDF of X!

$$F_X(a) = \begin{cases} 0 & a < 1 \\ \frac{1}{2} & 1 \le a < 2 \\ \frac{1}{2} + \frac{1}{3} = \frac{5}{6} & 2 \le a < 3 \\ \frac{1}{2} + \frac{1}{3} + \frac{1}{6} = 1 & a \ge 3 \end{cases}$$



- Even though X is a discrete R.V., CDF must be defined on real values.
- CDF is a step-function → discontinuity

Relation of PMF and CDF

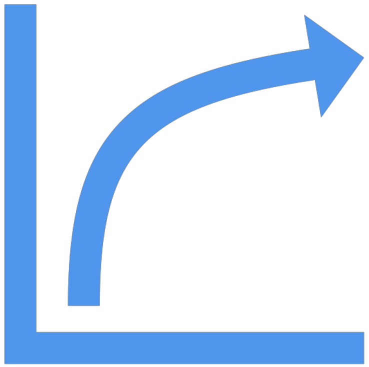
- For Discrete R.V., $\boldsymbol{F_X}(t)$ grows only by jumps in discrete steps

$$F_X(t) = F_X(t_i), \qquad for\ t_i \le t < t_{i+1}$$

$$F_X(t_{i+1}) = F_X(t_i) + P(X = t_{i+1})$$

- Hence, PMF can be obtained from CDF

$$P(X = x_k) = p_X(x_k) = F_X(x_k) - F_X(x_{k-1})$$



CONTINUOUS RANDOM VARIABLES

Probability Density Function (PDF) (1)

$$P(X \in B) = \int_B f_x(x)\, dx$$

- We say that X is a **continuous** random variable if there exists a non-negative function $f x(x)$, defined for all real $x \epsilon (-\infty, \infty)$, for any set B of real numbers
- The function $f x(x)$ is called the probability density function of random variable X.

*subscript x on $f_x(x)$ can be omitted if it's clear that $f(x)$ is PDF of random variable X

Probability Density Function (PDF) (2)

$$P(X \in B) = \int_B f_x(x)\,dx$$

- The probability that RV $\boldsymbol{X}$ will be in the interval $\boldsymbol{B},$ can be computed by taking the integral of the PDF $f_x(x)$ over $\boldsymbol{B}$

Probability Density Function (PDF) (3)

$$P(X \in B) = \int_B f_x(x)\, dx$$

- Letting B = [a, b], we obtain the probability within a range a ≤ x ≤ b

$$P(a \leq X \leq b) = \int_a^b f_x(x)\, dx$$

- This is an integral.

The required probability is the area under the curve f(x) between a and b !

Probability Density Function (PDF) (4)

- $f(x)$ must satisfy

$$1 = P(X \in (-\infty, \infty)) = \int_{-\infty}^{\infty} f_x(x)\, dx$$

- Which represents the probability of everything that may happen, or the entire sample space (S)

Probability Density Function (PDF) (5)

$$1 = P(X \in (-\infty, \infty)) = \int_{-\infty}^{\infty} f_x(x)\, dx$$

- Say we let a = b in the previous equation, then

$$P(X = a) = \int_{a}^{a} f_x(x)\, dx = 0$$

- this states that the probability that a continuous random variable will assume any particular value is zero.

Properties of PDF

- The following must hold for a valid PDF

1. Unlike PMF, since PDF is not a probability, the values of each function can be more than 1.

$$f_x(x) \geq 0, \qquad x \in R$$

2. Since continuous RV maps sample space into an uncountably infinite set of real numbers, then its probability has properties:

$$0 \leq \int_a^b f_X(x)dx \leq 1, \qquad \text{for } -\infty \leq a,b \leq \infty$$

$$\int_{-\infty}^{\infty} f_X(x)dx = 1$$

Relation between CDF and PDF

- Given the Cumulative Distribution Function F_x and Probability Density Function f_x
- We know

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

- or

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

Latihan

- Diketahui variabel acak (VA) $\boldsymbol{X}$ menunjukkan diameter dari sebuah lubang yang dibor pada komponen besi. Berdasarkan data yang sudah pernah dikumpulkan, distribusi dari $\boldsymbol{X}$ dapat dimodelkan sebagai sebuah PDF

$$f_X(x) = \begin{cases} 20e^{-20(x-12.5)} & x \geq 12.5 \\ 0 & lainnya \end{cases}$$

Apakah $f_x(x)$ PDF yang valid?

Periksa:

1. Apakah $f_x(x)$ selalu positif?
2. Apakah $\int_{-\infty}^{\infty} f_x(x)\,dx = 1$?

Latihan

- Diketahui variabel acak (VA) $\boldsymbol{X}$ menunjukkan diameter dari sebuah lubang yang dibor pada komponen besi. Berdasarkan data yang sudah pernah dikumpulkan, distribusi dari $\boldsymbol{X}$ dapat dimodelkan sebagai sebuah PDF

$$f_X(x) = \begin{cases} 20e^{-20(x-12.5)} & x \geq 12.5 \\ 0 & lainnya \end{cases}$$

1. Apakah $f_x(x)$ selalu positif?

   Iya $f_x(x) \geq 0, \quad x \geq 12.5$

2. Apakah $\int_{-\infty}^{\infty} f_x(x)\, dx = 1$?

   Iya $\int_{12.5}^{\infty} f_X(x)\, dx = \lim_{b \to \infty} \int_{12.5}^{b} 20e^{-20(x-12.5)}\, dx = \lim_{b \to \infty} \left(-e^{-20(x-12.5)} \Big|_{12.5}^{b} \right) = 1$

Latihan

- Diberikan VA kontinu X dengan PDF sebagai berikut

$$f_X(x) = \begin{cases} C(4x - 2x^2) & 0 < x < 2 \\ 0 & lainnya \end{cases}$$

a. Berapakah nilai C ?

b. Hitung P(X > 1) !

Latihan

a) Nilai **C** agar PDF valid:

Dapat ditunjukkan bahwa nilai **f(x)** $f(x) = C\left(4x - 2x^2\right) = C.x.\left(4 - 2x\right)$

≥ 0 pada interval 0 < x < 2, **bila** $C \geq 0$

Jadi, kita gunakan sifat ke-2 PDF untuk mencapatkan nilai C:

$$C\int_0^2 (4x - 2x^2)\, dx = C\left(2x^2 - \frac{2x^3}{3}\right)\Bigg|_0^2 = 1$$

$$hence, \quad C = \frac{3}{8}$$

b) $P(X > 1) = ?$ **Selesaikan problem (b) !**

Cumulative Distribution Function

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(y)\, dy \qquad x \in \Re$$

- Differentiating both sides yields:

$$f_X(x) = \frac{dF_X(x)}{dx}$$

- Other identities:

$$\bullet\ P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F_X(b) - F_X(a)$$

$$\bullet\ P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b)$$

The last one is ONLY for continuous case !

Latihan

Diberikan PDF sebuah VA kontinu **X**:

$$f_X(x) = \begin{cases} 20e^{-20(x-12.5)} & x \geq 12.5 \\ 0 & lainnya \end{cases}$$

(a) Tentukan CDF dari X ! (d.k.l, cari *F*(x)) !

(b) Hitunglah *P*(12.5 < X < 12.6) menggunakan *F*(x) !

- Sesuai definisi, $F(x) = P(X \le x)$

$$= \int_{-\infty}^{x} 20e^{-20(y-12.5)}\,dy$$

$$= \int_{12.5}^{x} 20e^{-20(y-12.5)}\,dy$$

$$= -e^{-20(y-12.5)}\Big|_{12.5}^{x} = 1 - e^{-20(x-12.5)}$$

Jadi, didapatkan **F(x)** :

$$F(x) = \begin{cases} 0 & x < 12.5 \\ 1 - e^{-20(x-12.5)} & x \ge 12.5 \end{cases}$$

Karena CDF harus didefinisikan pada **semua bilangan riil**

- Dan karena **X** VA kontinu $P(12.5 < X < 12.6)$

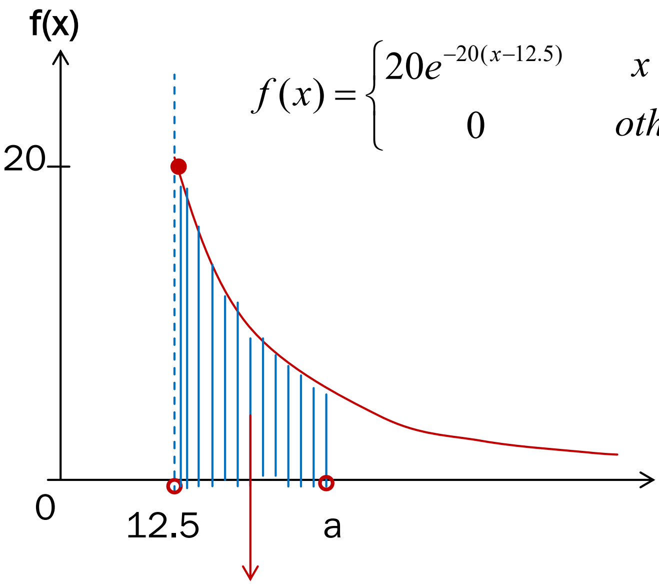
$$= P(12.5 < X \le 12.6)$$

$$= F(12.6) - F(12.5)$$

$$= 1 - e^{-20(12.6-12.5)} - (1 - e^{-20(12.5-12.5)})$$

$$= 1 - 0.135 = 0.865$$

Graf PDF

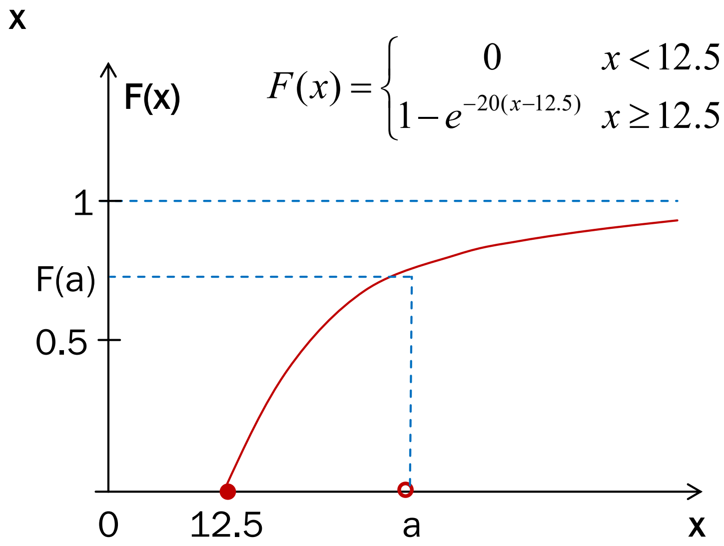


$$f(x) = \begin{cases} 20e^{-20(x-12.5)} & x \geq 12.5 \\ 0 & otherwise \end{cases}$$

This area is **F(a)**

$$F(a) = 1 - e^{-20(a-12.5)}$$

Graf CDF



$$F(x) = \begin{cases} 0 & x < 12.5 \\ 1 - e^{-20(x-12.5)} & x \geq 12.5 \end{cases}$$

What if We Have 2 or More Random Variables

- We are often interested in the relationships between two or more random variables.

- In an experiment into the possible causes of cancer, we might be interested in the relationship between
  - X: average number of cigarettes smoked daily
  - Y: age at which an individual contracts cancer

JOINTLY DISTRIBUTED RANDOM VARIABLES

Joint Probability Mass Function

- Assuming Discrete RVs, recall PMF of RV X

$$p(a_i) = P(X = a_i)$$

- Thus, we define the joint probability mass function of X and Y

$$p\left(a_i, b_j\right) = P(X = a_i, Y = b_j)$$

- Where for all possible values of X and Y

$$\sum_i \sum_j p\left(a_i, b_j\right) = 1$$

Joint Probability Mass Function (2)

- Given the joint PMF of RV X and Y

$$p(a_i, b_j) = P(X = a_i, Y = b_j)$$

- The PMF for discrete RV $X = x_i$, RV Y can have any value...

$$\{X = a_i\} = \bigcup_j \{X = a_i, Y = b_j\}$$

- Thus its probability

$$P(X = a_i) = P\left(\bigcup_j \{X = a_i, Y = b_j\}\right) = \sum_j P(X = a_i, Y = b_j) = \sum_j p(a_i, b_j)$$

Joint Probability Mass Function (3)

- So the PMF of both X and Y

$$
\begin{aligned}
P\{X = x_i\} &= \sum_j P\{X = x_i, Y = y_j\} \\
&= \sum_j p(x_i, y_j)
\end{aligned}
$$

$$
\begin{aligned}
P\{Y = y_j\} &= \sum_i P\{X = x_i, Y = y_j\} \\
&= \sum_i p(x_i, y_j)
\end{aligned}
$$

Marginal Probability Distribution

- For Discrete R.V.s
- Knowing $\boldsymbol{P}(\boldsymbol{X} = xi, \boldsymbol{Y} = yj)$ can determine $\boldsymbol{P}(\boldsymbol{X} = xi)$ & $\boldsymbol{P}(\boldsymbol{Y} = yj)$

- Obtained by summing the joint probability distribution over the values of the other random variable.

$$P(X = x_i) = \sum_j P(X = x_i, Y = y_j) = \sum_j p(x_i, x_j)$$

$$P(Y = y_i) = \sum_i P(X = x_i, Y = y_j) = \sum_i p(x_i, x_j)$$

Joint Cumulative Distribution

- If X and Y are random variables, the joint cumulative distribution function of X and Y is defined as:

$$F_{XY}(a, b) = F(X \leq a, Y \leq b)$$

- Read as: the probability that RV X is lower or equal to value x and *at the same time* RV Y is lower or equal to value y

Joint Cumulative Distribution

- If we know $F(x, y)$, we can theoretically compute the probability of any statement concerning the values of X and Y

$$F_{XY}(x, y) = P(X \leq x, Y \leq y)$$

- Logic:$P(X \leq x)$.... The value of Y doesn't matter......

$$\begin{aligned} F_X(x) &= P(X \leq x) \\ &= P(X \leq x, Y < \infty) = F(x, \infty) \end{aligned}$$

- $P(Y \leq y)$.... The value of X doesn't matter......

$$\begin{aligned} F_Y(y) &= P(Y \leq y) \\ &= P(X < \infty, Y \leq y) = F(\infty, y) \end{aligned}$$

Joint CDF for Discrete RVs

- For Discrete R.V.s
- By definition

$$F_{XY}(a,b) = P(X \leq a, Y \leq b)$$

- In discrete cases, it becomes

$$\begin{aligned} F_{XY}(a,b) &= P(X \leq a, Y \leq b) \\ &= \sum_{i:x_i \leq a} \sum_{j:y_j \leq b} P(X = x_i, Y = y_j) \\ &= \sum_{i:x_i \leq a} \sum_{j:y_j \leq b} p(x_i, y_j) \end{aligned}$$

Contoh

- Diketahui 3 baterai dipilih secara random dari kumpulan baterai berikut
  - Sekumpulan 3 baterai baru,
  - Sekumpulan 4 baterai lama tapi masih bisa digunakan, dan
  - Sekumpulan 5 baterai rusak.
- Bila VA X & Y didefinisikan sebagai berikut
  - X: jumlah baterai baru yang terpilih
  - Y: jumlah baterai lama terpilih yang masih bisa digunakan

Tentukan *joint probability mass function* dari X dan Y $p(x, y)$ untuk semua kemungkinan......

PMF – Lihat semua kemungkinan...

- Joint PMF dari X dan Y, $p(i,j) = P(X = i, Y = j)$ adalah...

$i \backslash j$	0	1	2	3	Row Sum $= P\{X = i\}$
0	$\frac{10}{220}$	$\frac{40}{220}$	$\frac{30}{220}$	$\frac{4}{220}$	$\frac{84}{220}$
1	$\frac{30}{220}$	$\frac{60}{220}$	$\frac{18}{220}$	0	$\frac{108}{220}$
2	$\frac{15}{220}$	$\frac{12}{220}$	0	0	$\frac{27}{220}$
3	$\frac{1}{220}$	0	0	0	$\frac{1}{220}$
Column Sums $= P\{Y = j\}$	$\frac{56}{220}$	$\frac{112}{220}$	$\frac{48}{220}$	$\frac{4}{220}$	

(a) Hitung ***F***(1, 1) !

(b) Hitung ***P***(**Y** = 2) !

Joint Probability Density Function

- For continuous R.V.s
- For every set C of pairs of real numbers

$$P((X,Y) \in C) = \iint_{(x,y)\in C} f(x,y)\, dx\, dy$$

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

Joint Probability Density Function (2)

- For continuous R.V.s
- We have the joint probability density function

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

- R.V. X and Y are jointly continuous if there exists a function f(x, y), having the property:

$$f_{XY}(x,y) \ge 0$$

and

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

Joint Cumulative Distribution Function

- The CDF of RV X… $P(X \in A)$ while Y can take any value….
- Hence

$$P(X \in A) = P(X \in A, Y \in (-\infty, \infty))$$
$$= \int_A \int_{-\infty}^{\infty} f(x, y)\, dy\, dx$$
$$= \int_A f_X(x)\, dx$$

Joint Cumulative Distribution Function

- For continuous R.V.s
- By definition

$$F_{XY}(a,b) = P(X \le a, Y \le b)$$

- In continuous case, it becomes

$$\begin{aligned} F_{XY}(a,b) &= P(X \le a, Y \le b) \\ &= P(X \in (-\infty, a], Y \in (-\infty, b]) \\ &= \int_{-\infty}^{b}\int_{-\infty}^{a} f(x,y)\, dx\, dy \end{aligned}$$

Marginal Probability Distribution

- For continuous R.V.s
- Obtained by integrating the joint probability distribution over the values of the other random variable.

$$P(X \in A) = P(X \in A, Y \in (-\infty, \infty))$$

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

$$= \int_A f_X(x)\, dx$$

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

- Similarly,

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

Latihan

- *Joint density function* dari **X** dan **Y** adalah

$$f(x,y) = \begin{cases} 2e^{-x}e^{-2y} & 0 < x < \infty, 0 < y < \infty \\ 0 & lainnya \end{cases}$$

(a) Hitung $P(X > 1, Y < 1)$

(b) $\boldsymbol{F_X}(a) = ?$

Latihan

- *Joint density function* dari **X** dan **Y** adalah

$$f(x,y) = \begin{cases} 2e^{-x}e^{-2y} & 0 < x < \infty, 0 < y < \infty \\ 0 & lainnya \end{cases}$$

(a) Hitung $P(X > 1, Y < 1)$

$$P(X > 1, Y < 1) = \int_0^1 \int_1^\infty 2e^{-x}e^{-2y}\ dx\ dy = e^{-1}(1 - e^{-2})$$

(b) $\boldsymbol{F_X}(a) = ?$

$$F_X(a) = P(X < a) = P(X < a, Y > 0) = \int_0^a \int_0^\infty 2e^{-x}e^{-2y}\ dy\ dx$$

Joint Probability Distribution For n R.V.s

$$F(a_1, a_2, \ldots, a_n) = P(X_1 \le a_1, X_2 \le a_2, \ldots, X_n \le a_n)$$

- In discrete case:

$$p(x_1, x_2, \ldots, x_n) = P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$$

- In continuous case:

$$P(X_1 \in A_1, X_2 \in A_2, \ldots, X_n \in A_n)$$

$$= \int_{A_n} \int_{A_{n-1}} \ldots \int_{A_1} f(x_1, \ldots, x_n)\, dx_1\, dx_2 \ldots dx_n$$

INDEPENDENT RANDOM VARIABLES

Independent Random Variables

- The random variables X and Y are said to be independent if and only if

- If discrete: $$p(x, y) = p_X(x)p_Y(y) = P(X = x).P(Y = y)$$

- If continuous: $$f(x, y) = f_X(x)f_Y(y)$$

N Independent Random Variables

- If there are ***n*** *independent* random variables ...

$$P(X_1 \in A_1, X_2 \in A_2, \ldots, X_n \in A_n) = \prod_{i=1}^{n} P(X_i \in A_i)$$

N Independent Random Variables (2)

- If there are ***n*** *independent* random variables ...

$$P(X_1 \in A_1, X_2 \in A_2, \ldots, X_n \in A_n) = \prod_{i=1}^{n} P(X_i \in A_i)$$

- As a consequence:
  - Joint cumulative distribution function of X and Y

$$F(a,b) = F_X(a)F_Y(b)$$

  - And, we have (for discrete & continuous case)

$$P(X \in A, Y \in B) = P(X \in A).P(Y \in B)$$

CONDITIONAL DISTRIBUTION

Conditional Distributions

- Discrete RV
- If X and Y are discrete random variables, we define the conditional probability mass function of X given that Y = y, by

$$\begin{aligned} p_{X|Y}(x \mid y) &= P(X = x \mid Y = y) \\ &= \frac{P(X = x, Y = y)}{P(Y = y)} \\ &= \frac{p(x, y)}{p_Y(y)} \end{aligned} \qquad p_Y(y) > 0$$

Conditional Distributions (2)

- For a Continuous RV
- If X and Y have a joint probability density function f(x, y), then the conditional probability density function of X, given that Y = y is

$$f_{X|Y}(x \mid y) = \frac{f(x, y)}{f_Y(y)} \qquad f_Y(y) > 0$$

- So,

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

Latihan 1

- Diketahui $p(x, y)$ adalah *joint probability mass function* dari X dan Y, dengan keterangan:

$$p(0,0) = 0.4 \qquad p(0,1) = 0.2 \qquad p(1,0) = 0.1 \qquad p(1,1) = 0.3$$

- Hitung *conditional probability mass function* dari X bila diketahui Y = 1 !

Latihan 2

- Diketahui *joint density* dari X dan Y adalah

$$f(x,y)=\begin{cases}\frac{12}{5}x(2-x-y) & 0<x<1, 0<y<1\\ 0 & lainnya\end{cases}$$

- Hitung *conditional density* dari X, bila diketahui Y = y, di mana 0 < y < 1 !