Reminder of Random Variables II

Dr Ahmad Khonsari
ECE Dept.
The University of Tehran

5. System of RVs: jointly distributed RVs

Basic notes:

- sometimes it is required to investigate two or more RVs;

- we assume that RVs X and Y are defined on some probability space.

- Capital letters (i.e. X, Y) are random variables

and small letters (i.e. x, y are given constants)

5. System of RVs: jointly distributed RVs

Definition: joint probability distribution function (JPDF) of RVs X and Y is:

$$F_{XY}\ (x, y) = Pr\{X \leq x, Y \leq y\} \qquad (78)$$

For continuous RV., **Let us define**:

$$F_X\ (x) = Pr\{X \leq x\} \quad F_Y\ (y) = Pr\{Y \leq y\} \quad x, y \in R, \qquad (79)$$

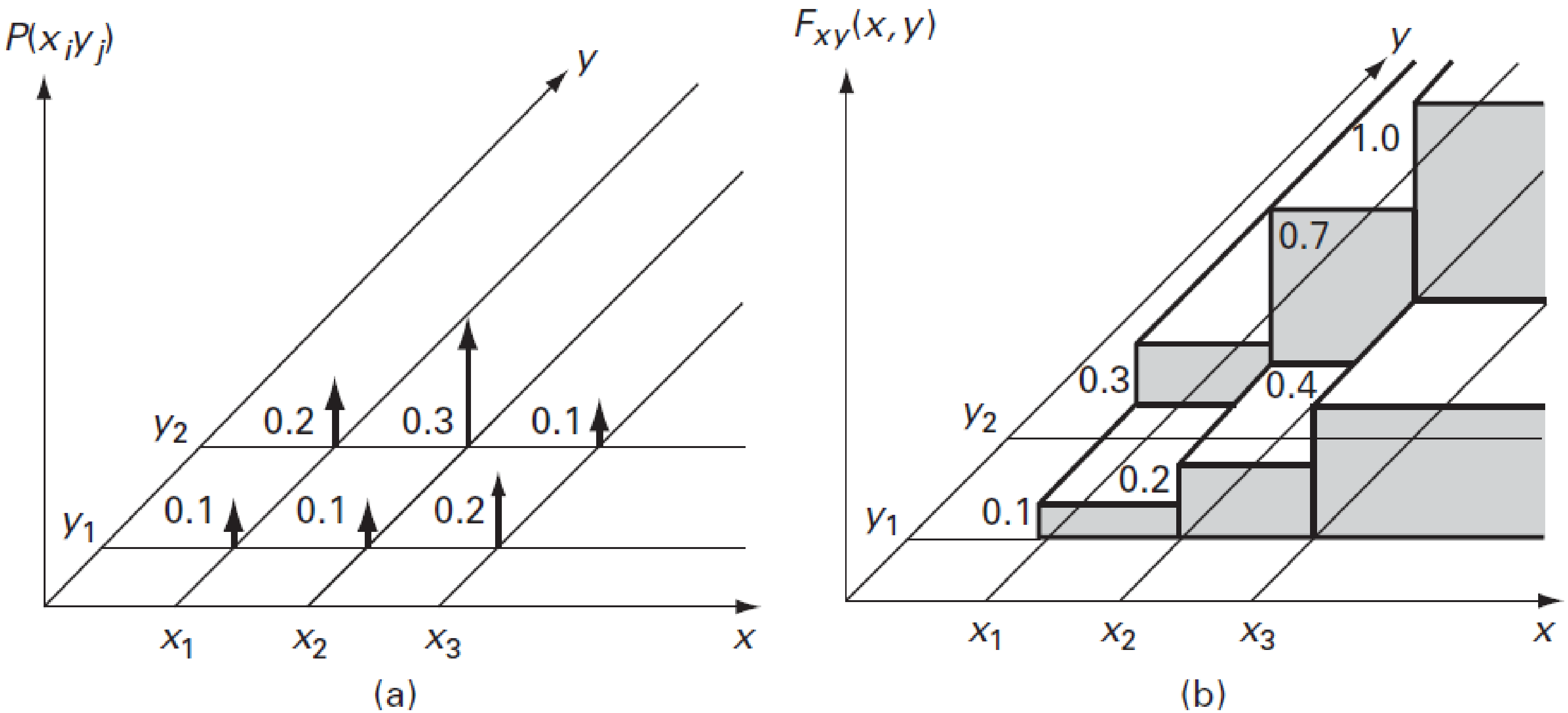
$F_X\ (x)$and $F_Y\ (y)$ are called marginal PDFs.

Marginal PDF can be derived form JPDF:

marginalize=neutralize=summing up to 1

$$F_X\ (x) = \lim_{y \to \infty} F_{XY}\ (x, y) = F_{XY}\ (x, \infty) \qquad (80)$$

$$F_Y\ (y) = \lim_{x \to \infty} F_{XY}\ (x, y) = F_{XY}\ (\infty, y)$$



(a) The joint probability distribution and
(b) the joint distribution function.

Definition: if $F_{XY}\ (x, y)$ is differentiable then the following function:

$$f_{XY}(x, y) = \frac{d^2}{dxdy} F_{XY}\ (x, y) \qquad (81)$$

$$= Pr\{x \le X \le x + dx, y \le Y \le y + dy\}$$

is called joint probability density function (jpdf).

Assume then that X and Y are discrete RVs.

Definition: joint probability mass function (Jpmf) of discrete RVs X and Y is:

$$f_{XY}(x,y)=\Pr\{X=x,Y=y\} \quad (82)$$

Let us define:

$$f_X(x)=\Pr\{X=x\} \qquad f_Y(y)=\Pr\{Y=y\} \quad (83)$$

- these functions are called marginal probability mass functions (Mpmf).

Marginal pmfs can be derived from Jpmf:

$$f_x(x)=\sum_{\forall y} f_{XY}(x,y), \qquad f_Y(y)=\sum_{\forall x} f_{XY}(x,y) \quad (84)$$

با داشتن تابع توزیع توأم (یا تابع توزیع احتمال توأم) می توان جرم تک تک مولفه ها را بدست آورد، از جمله تابع توزیع حاشیه ای. ولی برعکس این موضوع درست نیست.

به عبارت دیگر با داشتن $P(X = x_i)$ و $P(Y = y_j)$ نمی توان $P(X = x_i, Y = y_j)$ را بدست آورد، ولی برعکس آن ممکن است.

$$P(X = x_i) = \sum_j P(x_i, y_j)$$

البته اگر پیشامدها مستقل باشند، به راحتی توزیع توأم از روی حاصلضرب ۲ توزیع کناری بدست می آید.

مثال: ۳ نوع باطری داریم: { نو=۳، کارکرده=۴ و خراب=۵}و میخواهیم سه باطری انتخاب کنیم.

پیشامدها $\begin{cases} X = \text{باطری برداشته شده نو باشد} \\ Y = \text{باطری برداشته شده کارکرده باشد} \end{cases}$

$$P(i,j) = P(X = i, Y = j) = ?$$

سه باطری برمی داریم احتمال آنکه صفر باطری سالم و صفر باطری کارکرده باشد. (احتمال اینکه هر سه باطری خراب باشد.)

$$P(0,0) = \frac{\binom{5}{3}}{\binom{12}{3}} = \frac{10}{220}$$

یک باطری کارکرده و دو تای دیگر خراب باشد. (صفر باطری سالم)

$$P(0,1) = \frac{\binom{4}{1}\binom{5}{2}}{\binom{12}{3}} = \frac{40}{220}$$

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

pmf متغیر x با جمع سطری و pmf متغیر y با جمع ستونی بدست می آیدو چون این اطلاعات از روی حاشیه ها (کناره ها) جدول بدست می آید، به آن ها توزیع های حاشیه ای x وy می گویند.

نکته ۱: $P(X|Y=y)$ توزیع احتمال است.

مثالی از احتمال شرطی:

$$\sum_x P(X|Y=2) = \frac{P(0,2)}{P(Y=2)} + \frac{P(1,2)}{P(Y=2)} + \frac{P(2,2)}{P(Y=2)} + \frac{P(3,2)}{P(Y=2)} = 1$$

$$= \frac{\frac{30}{220}}{\frac{48}{220}} + \frac{\frac{18}{220}}{\frac{48}{220}} + \frac{0}{\frac{48}{220}} + \frac{0}{\frac{48}{220}} == \frac{30}{48} + \frac{18}{48} = 1$$

پس $P(X|Y=y)$ توزیع احتمال است.

نکته ۲: $P(Y=2)$ یک احتمال است و توزیع احتمال نیست، چون مقدار آن $\frac{48}{220}$ است.

نکته ۳: توزیع های حاشیه ای یک خلاصه ای از یک توزیع توأم است.

5.1. Conditional distributions and Mean (we saw Cond. Prob. Before)

Discret RV **Definition:** the following expression:

$$Pr_{X|Y}\{.,y\} = Pr_{X|Y}\{.|y\} = f_{X|Y}(.,y) = f_{X|Y}(.|y) = \frac{\Pr\{X = \forall, Y = y\}}{\Pr\{Y = y\}} \quad (85)$$

- gives conditional PF of discrete RV X given that Y = y.

Conditional mean of RV X given Y = y can be obtained as:

$$E[X|Y = y] = \sum_{\forall i} x_i \, Pr_{X|Y}\{x|y\} \quad (86)$$

Continous RV **Definition:** the following expression:

$$f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)} \quad , \quad (87)$$

- gives conditional pdf of continuous RV X given that $Y = y$.

Conditional mean of RV X given Y = y from the following expression:

$$E[X|Y = y] = \int_{-\infty}^{\infty} x f_{X|Y} \, dx \quad (88)$$

5.1. Conditional distributions and Mean (we saw Cond. Prob. Before)

Conditional CDF:

$$F_{\mathrm{X|Y}}(x|y) = Pr(X \leq x|Y \leq y) = \frac{\Pr\{X \leq x, Y \leq y\}}{\Pr\{Y \leq y\}} = \frac{F_{\mathrm{X,Y}}(x,y)}{F_{\mathrm{Y}}(y)}$$

Conditional pdf:

$$f_{\mathrm{X|Y}}(x|y) = \lim_{\Delta y \to 0} f_X(x|Y \approx y) = \lim_{\Delta y \to 0} \frac{\partial}{\partial x} F_X(x|Y \approx y) = \frac{f_{\mathrm{X,Y}}(x,y)}{f_{\mathrm{Y}}(y)}$$

Note:

$$f_{\mathrm{X|Y}}(x|y) \neq \frac{\partial}{\partial x} F_X(x|y)$$

Since the condition in pdf is Y=y and the condition in cdf is $Y \leq y$

5.1. Conditional distributions and Mean (we saw Cond. Prob. Before)

Mixture Distribution:(page 239 Trivedi 1st ed.)

Conditoional density (pmf) can be extended to the case where X is discrete RV and Y is continuous RV (or vice versa)

5.2. Dependence and independence of RVs

Recall the definition of independent events E and F: P(EF)=P(E)P(F)
Definition: it is necessary and sufficient for two RVs X and Y to be independent:

$$F_{XY}(x,y) = F_X(x)F_Y(y) \text{ for all } x,y \quad (89)$$

- $F_{XY}(x,y)$ is the JPDF(=JCDF);
- $F_X(x)$ and $F_Y(y)$ are PDFs (CDFs) of RV X and Y .

Definition: it is necessary and sufficient for two continuous RVs X and Y to be independent:

$$f_{XY}(x,y) = f_X(x)f_Y(y) \text{ for all } x,y \quad (90)$$

- $f_{XY}(x,y)$ is the jpdf;
- $f_X(x)$ and $f_Y(y)$ are pdfs of RV X and Y .

Definition: it is necessary and sufficient for two discrete RVs X and Y to be independent:

$$p_{XY}(x,y) = p_{XY}(X = x, Y = \forall)p_Y(X = \forall, Y = y) \text{ for all } x,y \quad (91)$$

- $p_{XY}(x,y)$ is the Jpmf;
- $p_X(x)$and $p_Y(y)$ are pmfs (discrete RV) or pdfs (continuous RV)) of RV X and Y .

Let: D1, D2 be the outcomes of two rolls:
S=D1+D2. the sum of two rolls

Each roll of a 6-sided die is an independent trial,
D1,D2 are independent.
Are S ands D1 independent? No

1. p(D1=1,S=7)?
= p(D1=1)p(s=7)

2. p(D1=1,S=5)?
≠ p(D1=1)p(s=5)

Let: D1, D2 be the outcomes of two rolls:
S=D1+D2. the sum of two rolls

- Each roll of a 6-sided die is an independent trial,
- D1,D2 are independent.

Are S ands D1 independent?

1. p(D1=1,S=7)?
Event (S=7) : {(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)}

p(D1=1)p(S=7)=(1/6)(1/6)
=1/36 =p(D1=1,S=7)

2. p(D1=1,S=5)?
Event (S=5) : {(1,4), (2,3),(3,2),(4,1)}
p(D1=1)p(S=5)=(1/6)(4/36)
≠ 1/36=p(D1=1,S=5)

Independent events (D1=1),(S=7)

Dependent events (D1=1),(S=5)

All events (X=x,Y=y) must be independent for X,Y to be independent variables.

5.3. Measure of dependence

Sometimes RVs are not independent:

- as a measure of dependence correlation moment (covariance) is used.

Definition: covariance of two RVs X and Y is defined as follows:

$$\sigma_{XY} = K_{XY} = cov(X, Y) = E[(X - E[X])(Y - E[Y])] \tag{92}$$

- where from definition , we find that $K_{XY} = K_{YX}$.

One can find the covariance using the following formulas:

- assume that RV X and Y are **discrete**:

$$K_{XY} = \sum_{i} \sum_{j} (x_i - E[X])(y_j - E[Y\,])Pr\{X = x_i, Y = y_j\} \tag{93}$$

- assume that RV X and Y are **continuous**:

$$K_{XY} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_i - E[X])(y_i - E[Y\,])f_{XY}\,(x, y)dxdy \tag{94}$$

It is often easy to use the following expression :

$$\sigma_{XY} = K_{XY} = E[XY] - E[X]E[Y] \quad (95)$$

Problem with covariance: can be arbitrary in $(-\infty, \infty)$:

- problem: hard to compare dependence between different pair of RVs;
- solution: use correlation coefficient to measure the dependence between RVs.

Definition: correlation coefficient of RVs X and Y is defined as follows:

$$\rho_{XY} = \frac{K_{XY}}{\sigma[X]\sigma[Y]} = \frac{\sigma_{XY}}{\sigma[X]\sigma[Y]} \quad (96)$$

$$1 \leq \rho_{XY} \leq 1$$

- if $\rho_{XY} \neq 0$ then RVs X and Y are correlated and hence dependent;

- **Example:** assume we are given RVs X and Y such that Y = aX + b:

$$\begin{aligned} \rho_{XY} &= +1 \qquad a>0 \\ \rho_{XY} &= -1 \qquad a<0 \end{aligned} \quad (97)$$

Very important note:

- ρ_{XY} is the measure telling how close the dependence to **linear.**

Question: what conclusions can be made when $\rho_{XY} = 0$? They are uncorrelated

- or RVs X and Y are not LINEARLY dependent;
- when $\rho_{XY} = 0$ is does not mean that they are independent.

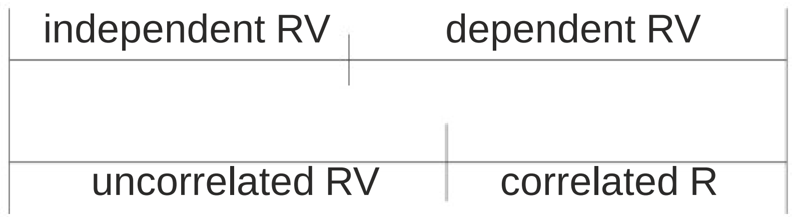


Fig: Independent and uncorrelated RVs.

What ρ_{XY} says to us:

- $\rho_{XY} \neq 0$: two RVs are correlated and also dependent;
- $\rho_{XY} = 0$: one can suggest that two RVs **MAY** BE independent;
- $\rho_{XY} = +1$ or $\rho_{XY} = -1$: RVs X and Y are linearly dependent.

5.4. (Expectations of product and Expectations of Sum) of correlated RVs

Mean:

- the mean of the product of two correlated RVs X,Y:

$$E[XY] = E[X]E[Y] + K_{XY} \tag{98}$$

- the mean of the product of two uncorrelated RVs X,Y:

$$E[XY] = E[X]E[Y] \tag{99}$$

Variance:

- the variance of the sum of two correlated RVs X,Y:

$$V[X+Y] = V[X] + V[Y] + 2K_{XY} \tag{100}$$

- the variance of the sum of two uncorrelated RVs X,Y:

$$V[X+Y] = V[X] + V[Y] \tag{101}$$

Now the Theory…

To capture this, define Covariance :

$$\boxed{\sigma_{XY} = E\{(X-\overline{X})(Y-\overline{Y})\}}$$

$$\boxed{\sigma_{XY} = \int\int (x-\overline{X})(y-\overline{Y})\, p_{XY}(x,y)\,dxdy}$$

If the RVs are both Zero-mean : $\boxed{\sigma_{XY} = \mathrm{E}\{XY\}}$

If X = Y: $\boxed{\sigma_{XY} = \sigma_X^2 = \sigma_Y^2}$

If X & Y are independent, then: $\boxed{\sigma_{XY} = 0}$

If $\sigma_{XY} = E\{(X - \overline{X})(Y - \overline{Y})\} = 0$

Say that X and Y are "uncorrelated"

If $\sigma_{XY} = E\{(X - \overline{X})(Y - \overline{Y})\} = 0$

Then $\underbrace{E\{XY\}}_{\text{Called "Correlation of X \&Y"}} = \overline{X}\,\overline{Y}$

So… RVs X and Y are said to be uncorrelated

if $E\{XY\} = E\{X\}E\{Y\}$

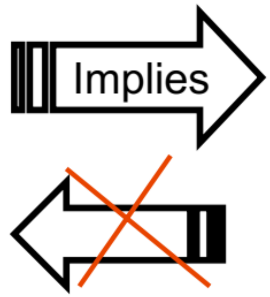
Independence vs. Uncorrelated

X & Y are Independent

$$f_{XY}(x, y)$$

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

PDFs Separate

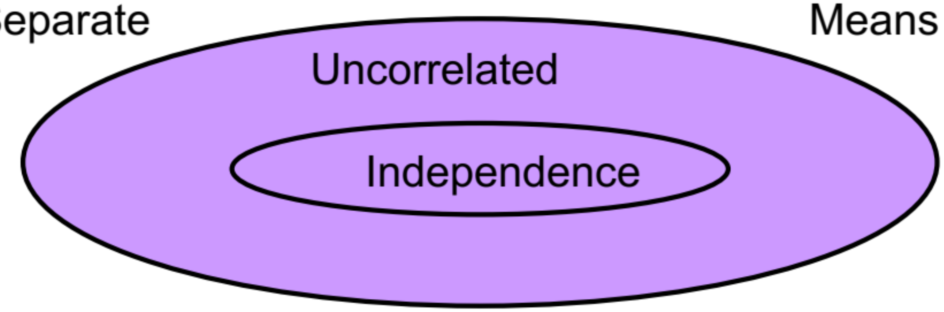


X & Y are Uncorrelated

$$E\{XY\}$$

$$= E\{X\}E\{Y\}$$

Means Separate



INDEPENDENCE IS A STRONGER CONDITION !!!!

Confusing Terminology…

Covariance : $\boxed{\sigma_{XY} = E\{(X-\bar{X})(Y-\bar{Y})\}}$

Correlation : $\boxed{E\{XY\}}$

Same if zero mean

Correlation Coefficient : $\boxed{\rho_{XY} = \dfrac{\sigma_{XY}}{\sigma_X \sigma_Y}}$

$\boxed{-1 \le \rho_{XY} \le 1}$

For Random Vectors…

$$\mathbf{x} = [X_1 \; X_1 \; \cdots \; X_N]^T$$

Correlation Matrix :

$$\boxed{\mathbf{R}_\mathbf{x} = E\{\mathbf{x}\mathbf{x}^T\} = \begin{bmatrix} E\{X_1X_1\} & E\{X_1X_2\} & \cdots & E\{X_1X_N\} \\ E\{X_2X_1\} & E\{X_2X_2\} & \cdots & E\{X_2X_N\} \\ \vdots & \vdots & \ddots & \vdots \\ E\{X_NX_1\} & E\{X_NX_2\} & \cdots & E\{X_NX_N\} \end{bmatrix}}$$

Covariance Matrix :

$$\boxed{\mathbf{C}_\mathbf{x} = E\{(\mathbf{x} - \overline{\mathbf{x}})(\mathbf{x} - \overline{\mathbf{x}})^T\}}$$

ارتباط شکل های 1 تا 4 را با روابطA,B,C,D را مشخص نمایید.

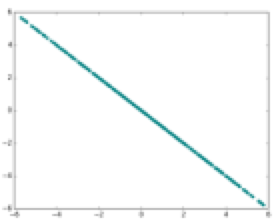
A. $\rho(X,Y) = 1$

B. $\rho(X,Y) = -1$

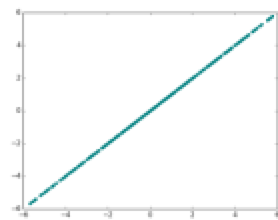
C. $\rho(X,Y) = 0$

D. Other

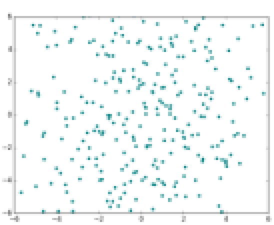
1.



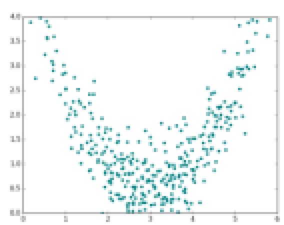
2.



3.



4.



1-B $Y = -\frac{\sigma_Y}{\sigma_X}X + b$

خطی با ضریب زاویه منفی است . به مقدار ضریب زاویه هم توجه کنید و سعی کنید ان را متوجه شوید

2-A $Y = \frac{\sigma_Y}{\sigma_X}X + b$

خطی با ضریب زاویه مثبت است . به مقدار ضریب زاویه هم توجه کنید و سعی کنید ان را متوجه شوید

3- C. $\rho(X,Y)$ =0 (ناهمبسته)

4- C. $\rho(X,Y)$ =0 $Y = X^2$

همانطور که تاکید شد ***همبستگی*** " خطی بودن " را اندازه می گیردو شکل 4 نشان میدهد با انکه کوواریانس XوY صفر می باشد . این دو متغیر به طور "غیر خطی " با هم رابطه دارند

6. Pdf of Sum of independent RVs

We consider independent RVs X and Y with probability functions:

$$P_X(x) = \Pr\{X = x\}, P_Y(y) = \Pr\{Y = y\} \qquad (102)$$

PMF of RV Z, Z = X + Y is defined as follows (i.e. convolution operation.)

$$\Pr\{Z = z\} = \sum_{k=-\infty} \Pr\{X = k\} \Pr\{Y = z - k\} \qquad (103)$$

- if $X = k$, then, Z take on z $(Z = z)$ if and only if $Y = z - k$.

If RVs X and Y are continuous:

$$f_X(x) \odot f_Y(y) = \int_{-\infty}^{\infty} f_X(z - y) f_Y(y) dy = \int_{-\infty}^{\infty} f_Y(z - x) f_X(x) dx \qquad (104)$$

Exercise: CDF of sum of 2 independent RVs :$\boldsymbol{F_z(z) = F_x(z) \odot f_y(z) = f_x(z) \odot F_y(z)}$

Q: what is pdf of the sum of two RVs generally

* سوال ① تابع چگالی احتمال $Z = X+Y$ را بدست آورید

(5-21) $F_Z(z) = \iint I(X+Y \le z)\, f_{XY}(x,y)\,dx\,dy = \int_{-\infty}^{\infty}\left[\int_{-\infty}^{z-y} f_{XY}(x,y)\,dx\right]dy$

$f_Z(z) = \int_{-\infty}^{\infty}\left[\frac{\partial}{\partial z}\int_{-\infty}^{z-y} f_{XY}(x,y)\,dx\right]dy$ مشتق نسبت به z می‌گیریم.

(مشتق لایب‌نیتز) مشتق انتگرال یعنی وقتی حدود انتگرال تابعی از متغیر دیفرانسیل هستند

(5-94) 133 $\frac{d}{dz}\int_{a(z)}^{b(z)} h(z,y)\,dy = h(z,b(z))\,b'(z) - h(z,a(z))\,a'(z) + \int_{a(z)}^{b(z)} \frac{\partial}{\partial z} h(z,y)\,dy$

[illegible]

پس $\frac{\partial}{\partial z}\int_{-\infty}^{z-y} f_{XY}(x,y)\,dx = f_{XY}(z-y,y)\times 1 - f_{XY}(-\infty,y)\times 0 + \int_{-\infty}^{z-y}\frac{\partial f_{XY}(x,y)}{\partial z}\,dx$

مشتق x,y نسبت به z صفر است

$= f_{XY}(z-y,y)$

صفحه ۱۱۵ کتاب پاپولیس

پس $f_Z(z) = \int_{-\infty}^{\infty} f_{XY}(z-y,y)\,dy$

[illegible]

در صورت استقلال X,Y داریم: $f_{XY}(x,y)=f_X(x)f_Y(y)$

$$f_Z(z)=\int_{-\infty}^{\infty} f_X(z-y)f_Y(y)\,dy=\int_{-\infty}^{\infty} f_X(x)f_Y(z-x)\,dx=f_X(z)\circledast f_Y(z)$$

برای حالت استقلال CDF ، Z برحسب CDF های X و Y قابل بدست آوردن است

(3.21) را به کمک انتگرال استیلتیز بازنویسی کنیم:

$$F_Z(z)=\iint I(x+y\le z)\,dF_X(x)\,dF_Y(y)=\int_{-\infty}^{\infty}\int_{-\infty}^{z-y} dF_X(x)\,dF_Y(y)$$

پس
$$F_Z(z)=\int_{-\infty}^{\infty}F_X(z-y)\,dF_Y(y)=\int_{-\infty}^{\infty}F_Y(z-x)\,dF_X(x)$$

ملاحظه کنید در دو فرم به تقارن نقش X و Y است.

به جای نوشتن $dF_Y(y)$ ، $f_Y(y)dy$ و $dF_X(x)$ ، $f_X(x)dx$

شکل کانولوشن برای CDF $F_Z(z)$ را داریم:

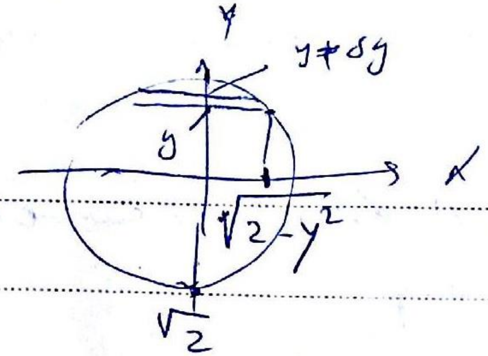
$$F_Z(z)=F_X(z)\circledast f_Y(z)=f_X(z)\circledast F_Y(z)$$

• سوال (۳۰) $z = x^2 + y^2$ ، اگر چگالی توأم $f_{XY}(x,y)$ موجود باشد ، تابع توزیع احتمال Z را بدست آورید.

$$F_Z(z) = P(X^2+Y^2 \leq z) = \iint I(x^2+y^2 \leq z)\, f_{XY}(x,y)\,dx\,dy$$

ناحیه $\{(x,y);\, x^2+y^2 \leq z\}$ دایره‌ای به شعاع $\sqrt{z}$ حول مبدأ است.

پس: $$F_Z(z) = \int_{-\sqrt{z}}^{\sqrt{z}} \left[\int_{-\sqrt{z-y^2}}^{\sqrt{z-y^2}} f_{XY}(x,y)\,dx \right] dy$$



$$\frac{\partial}{\partial z}F_Z = \frac{\partial}{\partial}\int_{-\sqrt{2}}^{\sqrt{2}}[\;]\;) = \int_{-\sqrt{2}}^{\sqrt{2}}\frac{\partial}{\partial}\left(\int_{-\sqrt{2-y^2}}^{\sqrt{2-y^2}} dx\right)dy$$

از $F_Z(z)$ نسبت به z مشتق بگیریم و ترتیب انتگرال

دو را عوض می‌کنیم. داخل را که با قاعده لایب‌نیتز

مشتق بگیریم درون [ ] داریم.

$$f_{xy}(\sqrt{z-y^2},y)\,\frac{1}{2\sqrt{z-y^2}} -$$

$$f_{xy}(-\sqrt{z-y^2},y)\left(-\frac{1}{2\sqrt{z-y^2}}\right) + 0$$

پس

$$f_Z(z) = \frac{d}{dz}F_Z(z) = \int_{-\sqrt{z}}^{\sqrt{z}} \frac{1}{2\sqrt{z-y^2}}\Big[f_{xy}(\sqrt{z-y^2},y) +$$

$$f_{xy}(-\sqrt{z-y^2},y)\Big]\,dy$$

An interesting case that often arises in signal detection problems, and for which we have a closed-form solution, is when X and Y are independent normal variables with zero mean and common variance (Problem 5.13). □

P 118 Kobayashi

5.13 Independent normal distribution and exponential distribution. Let X_1 and X_2 be independent normal variables with zero mean and common variance σ^2. Show that $Z = X_1^2 + X_2^2$ is exponentially distributed with mean $2\sigma^2$:

$$f_Z(z) = \frac{1}{2\sigma^2} e^{-z/2\sigma^2} u(z). \tag{5.104}$$

Example 5.5: $\boldsymbol{R = \sqrt{X^2 + Y^2}}$**.** Let us set $Z = R^2$ in the previous example. In the context of detecting a signal of the form $S(t) = X\cos(\omega t - \phi) + Y\sin(\omega t - \phi)$, the RV $R = \sqrt{X^2 + Y^2}$ represents the **envelope** of the signal, i.e., $S(t) = R\cos(\omega t - \theta)$, where $\theta - \phi = \tan^{-1}\frac{Y}{X}$.

The distribution function of R is given by

$$F_R(r) = \int_{-r}^{r}\left[\int_{-\sqrt{r^2-y^2}}^{\sqrt{r^2-y^2}} f_{XY}(x, y)\, dx\right] dy. \tag{5.35}$$

Differentiation of the expression inside the square brackets leads, using Leibniz's rule again, to the following expression:

$$f_{XY}(\sqrt{r^2 - y^2}, y)\frac{1}{2}\frac{2r}{\sqrt{r^2 - y^2}} - f_{XY}(-\sqrt{r^2 - y^2}, y)\left(-\frac{1}{2}\frac{2r}{\sqrt{r^2 - y^2}}\right) + 0. \tag{5.36}$$

Thus, we obtain

$$f_R(r) = \frac{dF_R(z)}{dr} = \int_{-r}^{r}\frac{r}{\sqrt{r^2 - y^2}}\left[f_{XY}(\sqrt{r^2 - y^2}, y) + f_{XY}(-\sqrt{r^2 - y^2}, y)\right] dy. \tag{5.37}$$

Again, an important and useful case is found when X and Y are independent normal variables with common variance (see Section 7.5.1). □

5.6* **Leibniz's rule.**[9] In deriving (5.23), we used a special case of Leibniz's rule for differentiation under the integral sign.

THEOREM 5.1 (Leibniz's rule). *The following rule holds for differentiation of a definite integral, when the integration limits are functions of the differential variable:*

$$\boxed{\frac{d}{dz}\int_{a(z)}^{b(z)} h(z,y)\,dy = h(z,b(z))b'(z) - h(z,a(z))a'(z) + \int_{a(z)}^{b(z)} \frac{\partial}{\partial z}h(z,y)\,dy.} \tag{5.94}$$

In particular, if h is a function of y only, the rule reduces to

$$\boxed{\frac{d}{dz}\int_{a(z)}^{b(z)} h(y)\,dy = h(b(z))b'(z) - h(a(z))a'(z).} \tag{5.95}$$

□

(a) Define

$$\int_{-\infty}^{y} h(x)\,dx \triangleq H(y).$$

Then prove (5.95).

(b) Define

$$\int_{-\infty}^{y} h(z,x)\,dx \triangleq H(z,y) \text{ and } \frac{\partial H(z,y)}{\partial y} \triangleq g(z,y).$$

Then prove (5.94).

(c) Alternative proof of (5.94). Consider a function $G(a, b, c)$, where a, b, and c stand for $a(z)$, $b(z)$, and $c(z)$ respectively. By applying the *chain rule* to the function G, we have

$$\frac{dG(a,b,c)}{dz} = \frac{\partial G}{\partial a}a'(z) + \frac{\partial G}{\partial b}b'(z) + \frac{\partial G}{\partial c}c'(z). \tag{5.96}$$

Consider a special case

$$c(z) = z \text{ and } G(a,b,c) \triangleq \int_a^b h(z,y)\,dy.$$

Then prove (5.94).

اگر از قاعده لایب نیتز، h فقط تابع y باشد داریم:

$$\frac{d}{dz}\int_{a(z)}^{b(z)} h(y)\,dy = h(b(z))\,b'(z) - h(a(z))\,a'(z) \qquad (5.95)$$

روش اثبات لایب نیتز:

الف) تعریف کنید $\int_{-\infty}^{y} h(x)\,dx \triangleq H(y)$

و از روی آن (5.95) را بدست آورید.

ب) تعریف کنید: $\int_{-\infty}^{y} h(z,x)\,dx \triangleq H(z,y)$

and

$$\frac{\partial H(z,y)}{\partial y} \triangleq g(z,y)$$

و سپس 5.94 با لایب نیتز حل را ثابت کنید.

2) روش دیگر اثبات 5.94: تابع $G(a,b,c)$ را در نظر بگیرید

که a و b و c به ترتیب $a(z)$ و $b(z)$ و $c(z)$ هستند. با اعمال قاعده

زنجیره‌ای روی تابع G داریم

$$\frac{dG(a,b,c)}{dz} = \frac{\partial G}{\partial a}a'(z) + \frac{\partial G}{\partial b}b'(z) + \frac{\partial G}{\partial c}c'(z) \qquad (5-96)$$

از حالت خاص $c(z)=z$ و $G(a,b,c) \triangleq \int_a^b h(z,y)\,dy$

اگر در رابطه قرار دهیم، رابطه 5.94 را نتیجه می‌گیرید

7. Indicator RVs

The **indicator random variable I{A}** associated with **event A** is defined as

$$\mathrm{I}\{A\}=\begin{cases}1 & if\ \ A\ occurs \\ 0 & if\ A\ does\ not\ occur\end{cases} \qquad (7.1)$$

Example: determine the expected number of heads in tossing a fair coin. **sample space** is

S={H,T}, with Pr{T}= Pr{H}= $\frac{1}{2}$.

Define the event *H* as the coin coming up heads, we define an indicator RV X_H associated with the **event *H***, such that:

X_H counts the number of heads obtained in this flip, i.e. it is 1 if the coin comes up heads and 0 , otherwise.

We write

$$X_H=\mathrm{I}\{H\}=\begin{cases}1 & if\ \ H\ occurs \\ 0 & if\ T\ occurs\end{cases}.$$

7. Indicator RVs

The expected number of heads obtained in one flip of the coin is simply the expected value of indicator variable X_H:

$$E[X_H] = E[I\{H\}]$$

$$= 1.\Pr\{H\} + 0.\Pr\{T\}$$

$$= 1.\left(\frac{1}{2}\right) + 0.\left(\frac{1}{2}\right) = \frac{1}{2}$$

Thus the expected number of heads obtained by one flip of a fair coin is 1/2.

Q: what is the difference between expected value and average case? Does make sense to define average with one flip ?

7. Indicator RVs

Lemma 7.1

Given a sample space S and an event A in the sample space S, let $X_A = I\{A\}$.

Then $$E[X_A] = Pr\{A\}$$

Proof:

By the definition of an indicator RV from equation (7.1) and the definition of expected value, we have

$$\begin{aligned} E[X_A] &= E[I\{A\}] \\ &= 1.\Pr\{A\} + 0.\Pr\{\bar{A}\} \\ &= \Pr\{A\} \end{aligned}$$

,where $\bar{A}$ denotes S - A, (i.e. the complement of A).

Thus the above lemma implies:

The expected value of an indicator RV associated with an event A is equal to the probability that A occurs.

7. Indicator RVs

Although indicator RVs may seem cumbersome for an application such as counting the expected number of heads on a flip of a single coin, they are useful for analyzing situations in which we perform repeated random trials.

Example: compute the expected number of heads in *n* tossing of a coin.

Let X denotes the total number of heads in the *n* coin flips, so that

$$X = \sum_{i=1}^{n} X_i$$

we take the expectation of both sides

$$\begin{aligned} E[X] &= E\left[\sum_{i=1}^{n} X_i\right] \\ &= \sum_{i=1}^{n} E[X_i] \\ &= \sum_{i=1}^{n} \frac{1}{2} \\ &= \frac{n}{2} \end{aligned}$$

7. Indicator RVs

We can compute the expectation of a random variable having a binomial distribution from equations

$$\binom{n}{k} = \frac{n}{k}\binom{n-1}{k-1}$$

and

$$\sum_{k=0}^{n} Bin(n-1; p) = 1.$$

7. Indicator RVs

Let X~Bin(n; p), q=1-p, By the definition of expectation, we have

$$E[X] = \sum_{k=0}^{n} k.\Pr\{X = x\}$$

$$= \sum_{k=0}^{n} k.Bin(n;p)$$

$$= \sum_{k=0}^{n} k\binom{n}{k} p^k q^{n-k} = \sum_{k=1}^{n} k\frac{n}{k}\binom{n-1}{k-1} p^k q^{n-k}$$

$$= np\sum_{k=1}^{n}\binom{n-1}{k-1} p^{k-1} q^{n-k} \qquad k-1=j=k$$

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

$$= np\sum_{k=0}^{n-1} Bin(n-1;p)$$

$$= np$$

7. Indicator RVs

Let X~Bin(n; p), q=1-p. Obtaining the same result using the linearity of expectation.

Let X_i denotes the number of successes in the *i* th trial. Then

$$E[X_i] = p.1 + q.0 = p$$

and by linearity of expectation, the expected number of successes for *n* trials is

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right]$$

$$= \sum_{i=1}^{n} E[X_i]$$

$$= \sum_{i=1}^{n} p$$

$$= np$$

7. Indicator RVs

Example: Let X~Bin(n; p), q=1-p calculate the variance of the distribution. Using $Var[X] = E[X^2] - E^2[X]$.,

we have $Var[X_i] = E\left[X_i^2\right] - E^2[X_i]$.

X_i only takes on the values 0 and 1, we have $X_i^2 = X_i$,

which implies $E\left[X_i^2\right] = E[X_i] = p$.

Hence, $Var[X_i] = p - p^2 = pq$

To compute the variance of X, we take advantage of the independence of the *n* trials; thus,

$$Var[X] = Var\left[\sum_{i=1}^{n} X_i\right]$$

$$= \sum_{i=1}^{n} Var[X_i]$$

$$= \sum_{i=1}^{n} pq$$

$$= npq$$

تفاوت نشانگر و برنولی ← برنولی متغیر تصادفی با حضور احتمال است و نشانگرها از تصادفی بودن بهره ندارد

← نشانگر ترکیب تصادفی بودن و احتمال حقیقی است، برای محاسبه شانس تصادفی بودن احتمال نشان داده می‌شود و ...

متغیر تصادفی نشانگر (Indicator RV): برای هر پیشامد A از فضای نمونه S را به زیرمجموعه

با نقطه اجتماع و شامل (collectively exhaustive ≠ هر نقطه را در بر بشود) افراز کرده

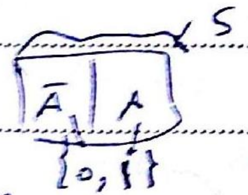
نشانگر پیشامد A را با متغیر تصادفی I_A نشان می‌دهیم و به صورت زیر تعریف می‌کنیم

$$I_A(s) = \begin{cases} 1 & \text{if } s \in A \\ 0 & \text{if } s \in \bar{A} \end{cases}$$

کاربرد متغیر تصادفی نشانگر: در مسائلی که دانش به تفصیل تابع توزیع وجود ندارد، به عنوان ابزاری با

استفاده از آن، محاسبات شمارش را انجام می‌دهیم

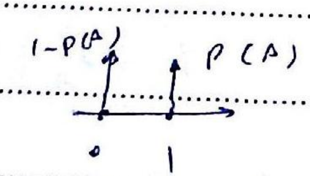
پس اگر A فقط وقتی رخ می‌دهد که $I_A = 1$ در مثال $\frac{2-16}{82}$ تمرین را با این ابزار حل کنید.



pmf (تابع جرم احتمال) متغیر تصادفی I_A (چون متغیر تصادفی دارای تابع [illegible] است) بصورت

زیر می‌باشد

$$P_{I_A}(0) = P(\bar{A}) = 1 - P(A)$$



$$P_{I_A}(1) = P(A)$$

https://www.statlect.com/fundamentals-of-probability/indicator-functions

Appendix: General Case: Let $X_1, X_2, \ldots X_k$ be continuous random variables

i. Their joint **Cumulative Distribution Function**, $F(x_1, x_2, \ldots x_k)$ defines the probability that simultaneously X_1 is less than x_1, X_2 is less than x_2, and so on; that is

$$F(x_1, x_2, \ldots, x_k) = P(X_1 < x_1 \cap X_2 < x_2 \cap \cdots X_k < x_k)$$

i. The cumulative distribution functions $F_1(x_1), F_2(x_2), \ldots, F_k(x_k)$ of the individual random variables are called their **marginal distribution function**. For any i, $F_i(x_i)$ is the probability that the random variable X_i does not exceed the specific value x_i.

iii. The random variables are **independent** if and only if

$$F(x_1, x_2, \ldots, x_k) = F_1(x_1)F_2(x_2)\cdots F_k(x_k)$$

$$\textit{or}\ \text{equivalently}$$

$$f(x_1, x_2, \ldots, x_k) = f_1(x_1)\, f_2(x_2)\cdots f_k(x_k)$$