A function $X : (\Omega, \mathcal{A}) \longrightarrow (\mathbb{R}, \mathcal{B})$

is called a random variable

iff for $^{\forall} A \in \mathcal{B}$, $X^{-1}(A) \in \mathcal{A}$.

$$P_X(X \in A) \text{ for } ^{\forall} A \in \mathcal{B}$$

$$:= P(\{w \in \Omega : X(w) \in A\})$$

$$P_X(B) = \int_B f_X(x)\,dx$$

Definition. Let X be a random variable, the probability (cumulative) distribution function of X, F_X, is the real-valued function on $\mathbb{R}$ defined by $F_X(x) = P_X(X \le x)$

$$F(x) = \int_{-\infty}^{x} f_X(t)\,dt$$

Properties: ① F_X is monotone nondecreasing

② F_X is right continuous

③ $\lim_{x \to \infty} F_X(x) = 1$

④ $\lim_{x \to -\infty} F_X(x) = 0$

$x \to -\infty$

ex)

$$f_X(x) = \begin{cases} 1 & x \in [0,1] \\ 0 & \text{otherwise} \end{cases}$$

$\underline{x < 0}$

$$F_X(x) = P_X(X \le x)$$
$$= \int_{-\infty}^{x} f_X(t)\,dt$$
$$= \int_{-\infty}^{x} f_X(t)\,dt = \int_{-\infty}^{x} 0\,dt$$
$$= 0$$

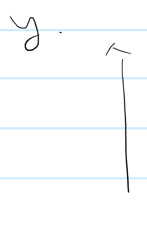
$x \in [0,1)$

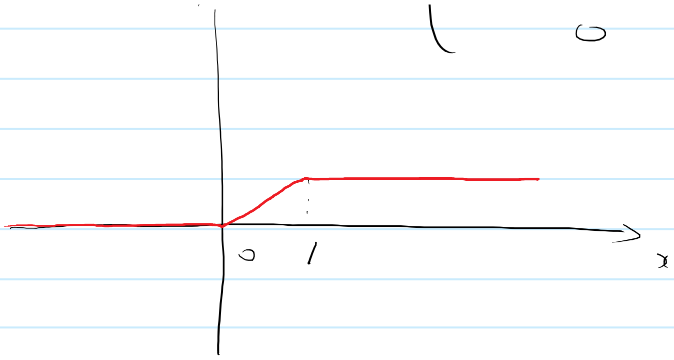
$$F_X(x) = P_X(X \le x)$$
$$= \int_{-\infty}^{x} f_X(t)\,dt$$
$$= \int_{-\infty}^{0} f_X(t)\,dt + \int_{0}^{x} f_X(t)\,dt$$
$$= 0 + \int_{0}^{x} 1\,dt$$
$$= x$$

$x > 1$

$$F_X(x) = P_X(X \le x)$$
$$= \underbrace{\int_{-\infty}^{0} f_X(t)\,dt}_{=0} + \underbrace{\int_{0}^{1} f_X(t)\,dt}_{=1} + \underbrace{\int_{1}^{x} f_X(t)\,dt}_{=0}$$
$$= 1$$

$$F_X(x) = \begin{cases} 1 & x > 1 \\ x & 0 \le x \le 1 \\ 0 & x < 0 \end{cases}$$





Note $P_X = P_Y$ $\qquad f(x) = g(x) \quad \forall x \in \mathbb{U}$

$\iff F_X = F_Y$

$P_X(X \le x) = P_Y(Y \le x)$

$\{\omega \in \Omega : X(\omega) \in (-\infty, x]\}$

$F_X \qquad (-\infty, x]$

$(a, b) \qquad [a, b]$

$(a, b]$

$[b, a)$

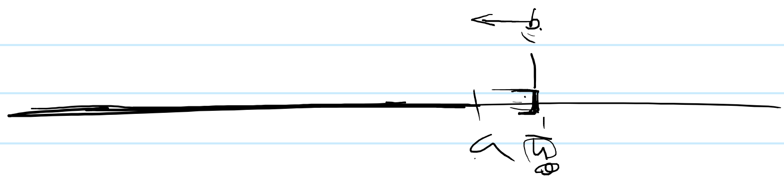
F

$(-\infty, a]$

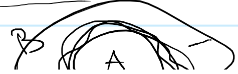
$P_X((-\infty, a)) = F_X$

$(-\infty, a]$

$$P_X((-\infty, a)) = \lim_{n \to \infty} P_X\left(\left(-\infty, a + \frac{1}{n}\right]\right)$$

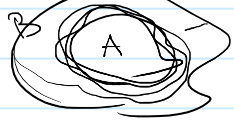


$P(A) \le P(B)$ if $A \subset B$



•furltiouX 1^, 쨋(10, 蜜 kmeda빠V억 sk페이지 3

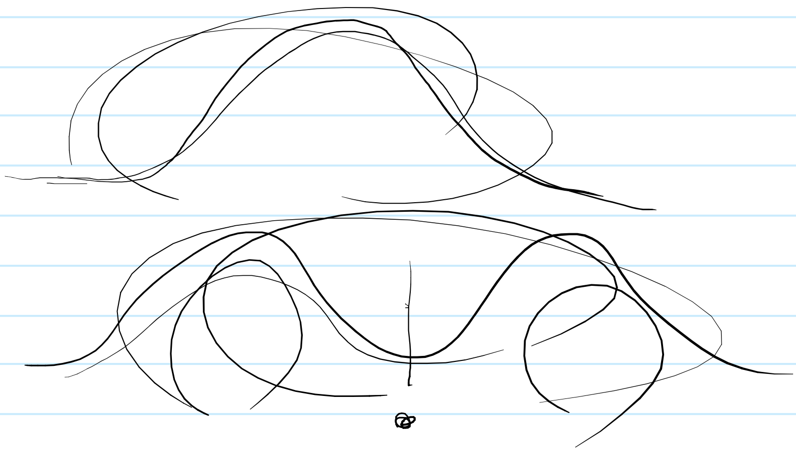
$P(A) \le P(B)$ if $A \subset B$



(Again, the classification of random variables,

X is discrete iff F_X is discrete. On the contrary,

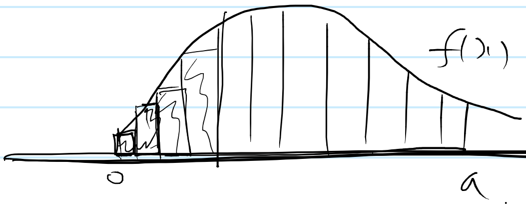
X is continuous iff F_X is continuous.)



<u>Definition</u> Let X be a random variable and g be a function. Then, the expectation of g(X), denoted E[g(X)], is defined by

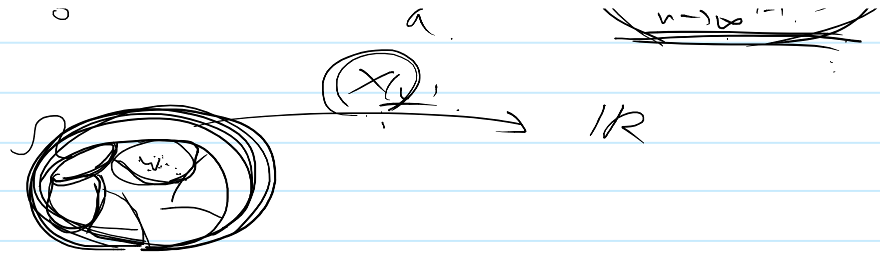
$$E[g(X)] = \int_{\Omega} g(X)\, dP$$

provided the integral exists.



$$\int_0^a f(x)\,dx = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i)\frac{1}{n}$$

(x_i)



Note $E[X] := \int_\Omega X \, dP$

mean, central tendency, μ

$$= \sum_x x f_X(x)$$

or

$$= \int_{\mathbb{R}} x f_X(x) \, dx$$

Theorem a. $E[a g_1(X) + b g_2(X) + c]$

$$= aE[g_1(X)] + bE[g_2(X)] + c$$

b. If $g_1(x) \geq 0$ for all x, then $E[g_1(X)] \geq 0$

c. If $g_1(x) \geq g_2(x)$ for all x, then $E[g_1(X)] \geq E[g_2(X)]$

d. If $a \leq g_1(x) \leq b$ for all x, then $a \leq E[g_1(X)] \leq b$

Definition Let X be a random variable with finite expectation, then, we define the variance of X, denoted by $Var(X)$, by

$$\sigma^2 = \underline{Var(X)} = E[(X - \overset{\mu}{E[X]})^2]$$

$$g(x) = (x - \mu)^2$$

Theorem

a. $V(aX+b) = a^2 \operatorname{Var} X$

b. $\operatorname{Var}(X) = E[X^2] - \mu^2$, $\mu = E[X]$

$$Y = g(X)$$

The Distribution of A function of A Random Variable.

Suppose we know the distribution of a given X and want to find the distribution of $g(X)$ for some function g.

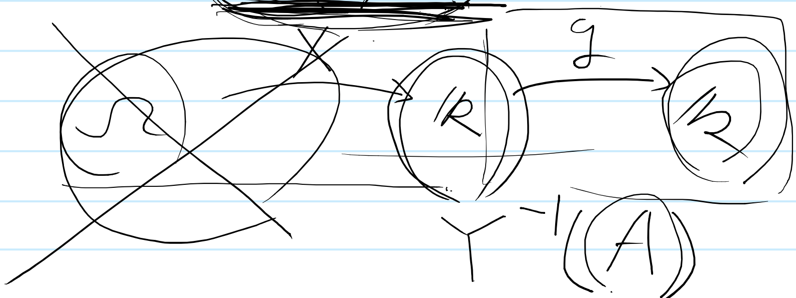
ex) Let X be uniformly distri-buted on $(0,1)$,

$$f_X(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

$$F_X(x) = \begin{cases} 0 & -\infty < x \le 0 \\ x & 0 < x < 1 \\ 1 & x \ge 1 \end{cases}$$

What is the distribution of $Y = X^n$, $g(x) = x^n$

F_Y $\quad 0 \le y \le 1$



$Y^{-1}(A)$

$$F_Y(y) = P_g[\,Y \le y\,] = P_g = \{$$

$$= P_Y[X^n \le y]$$

$$\overset{?}{=} P_X[X \le y^{\frac{1}{n}}]$$

$$(\Omega, \mathcal{A}, P) \xrightarrow{X} (\mathbb{R}, \mathcal{B}, P_X) \xrightarrow{g} (\mathbb{R}, \mathcal{B}, P_g)$$

$g(x)$

$$\boxed{\begin{aligned} &P_g(g(x) \in A) \\ &= P(g(x)^{-1}(A)) \end{aligned}}$$

ex) Let X be uniformly distributed over $(0,1)$, that is,

$$f_X(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

$$F_X(x) = \int_{-\infty}^{x} f_X(x)\,dx = \begin{cases} 0 & -\infty < x \le 0 \\ x & 0 < x < 1 \\ 1 & x \ge 1 \end{cases}$$

what is the distribution of $Y = X^n$, $g(x) = x^n$

$g($

$g(x)$

$g(x) = \omega$

$P_X(A)$ $\quad F_Y(y) = P_X((-\infty, y])$

$$F_Y(y) := P_Y[Y \le y]$$

$$:= P[\{w \in \Omega : Y(w) \le y\}]$$

$$:= P[\{w \in \Omega : X^n(w) \le y\}]$$

$$= P[\{w \in \Omega : X(w) \le y^{\frac{1}{n}}\}]$$

$y = x^n$

g^{-1} g

$X^n(\omega) \le y$

$(\omega) \le y^{\frac{1}{n}}$

$X^n(\omega) \le y$

$X(\omega) \le y^{\frac{1}{n}}$

$$= P[\{w \in \Omega : X(\omega) \le y^{n}\}]$$

S.I strictly increasing

$x \le y \Rightarrow f(x) \le f(y)$

$x < y \Rightarrow f(x) < f(y)$

$P_X[X \in (-\infty, y^{\frac{1}{n}}]]$

$$:= P_X[X \le y^{\frac{1}{n}}]$$

$$:= F_X(y^{\frac{1}{n}})$$

$$F_Y(y) = \begin{cases} 0 & y < 0 \\ y^{\frac{1}{n}} & 0 \le y \le 1 \\ 1 & y > 1 \end{cases}$$

$\Omega = [0,1]$

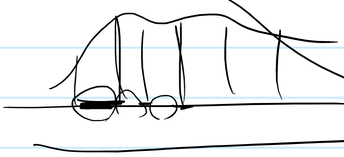
$(\Omega, \mathcal{A}, P)$ $\quad$ $(\mathbb{R}, \mathcal{B}, P_X)$

$X(\omega) = \omega$

$$F(x) = \int_{-\infty}^{x} f(t)\, dt$$

$$\frac{dF}{dx} = f(x)$$

$$\int_{\Omega} X(\omega)\, dP(\omega) = \int_{\mathbb{R}} x\, P_X(dx)$$



$$= \int_{\mathbb{R}} x f_X(x)\, dx$$

Theorem Let X be a continuous variable having probability density function f_X. Suppose that $g(x)$ is a strictly monotone (increasing or decreasing), differentiable

(and thus continuous)
function of x. Then the random variable Y defined by $Y = g(X)$ has a probability density function given by

$$f_Y(y) = \begin{cases} f_X\left(g^{-1}(y)\right) \left|\frac{d}{dy} g^{-1}(y)\right| & \text{if } y = g(x) \text{ for some } x \\ 0 & \text{if } y \neq g(x) \text{ for all } x \end{cases}$$

where $g^{-1}(y)$ is defined to equal that value of x such that $g(x) = y$.

Chapter 8

Ideal Models of Discrete Random Variables

① Bernoulli Random Variable

Bernoulli trial: random experiment having two mutually exclusive outcomes "Success" or "Fail" with $P(\text{Success}) = p$, $0 < p < 1$

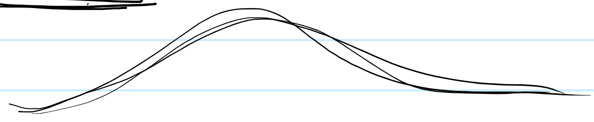
Define X by $X(\text{Success}) = 1$ and
$X(\text{Fail}) = 0$

Then, $P_X(X = 0) = 1 - p$
$P_X(X = 1) = p$

So, $f_X(x) = \begin{cases} 1-p & x = 0 \\ p & x = 1 \\ 0 & \text{otherwise} \end{cases}$

0 otherwise

$\boxed{B(p)}$



mean: $\mu = E[X] = 1 \cdot f_X(1) + 0 \cdot f_X(0)$

$= p$

variance: $\sigma^2 = E[X^2] - (E[X])^2$

$= 1^2 \cdot f_X(1) + 0^2 \cdot f_X(0) - p^2$

$= p - p^2 = p(1-p)$

(2) Binomial Random Variable

1. Each experiment consists of n independent "Bernoulli trials" with $P(\text{success}) = p$

2. Define X by the number of successes in the n trials

(ex) tossing a coin independently 100 times with $P(\text{Head}) = p$, $0 < p < 1$

40	17	5
	3	13
	0	0
	0	0.

that is, we do independent Bernoulli trials 100 times.

Then,

2^{100}

$\Omega = \{ \overbrace{HH \cdots H}^{100}, \overbrace{HH \cdots H}^{99}T, \ldots, \overbrace{TT \cdots T}^{100} \}$

ω_1

$\mathcal{S} = 2^{\Omega}$

$P(\omega_i) = ?$ $\quad i = 1, \ldots, 2^{100}$

$P(HH \cdots HT) = P(A_1 \cap \cdots \cap A_{100}^c)$

$A_i = \{ i\text{th trial returns } H \}$

9/10

<풀이

$A_i = \{ \text{ith trial returns } H \}$

$i = 1, \dots, 100$

$P(A \cap B) = P(A)P(B)$

$= P(A_1) \cdots P(A_{100}^c)$

$P(A_i) = p \quad P(A_i^c) = 1-p$

$P(\omega_i) = p^x (1-p)^{100-x}$, $x = \#$ of H in ω_i

Define X by $X(\omega) = \#$ of H. A

$$P_X[X = x] := P[\{\omega \in \Omega : X(\omega) = x\}]$$

$$= \sum_{\omega \in A} P(\omega) = \sum_{\omega \in A} p^{x_\omega}(1-p)^{100-x_\omega}$$

Note $k = x_\omega$, $n = 100$ and

Our case is unordered without replacement

$\frac{n!}{(n-x)!\,x!}$ Thus, the number is $\binom{n}{x}$

Therefore, $P_X[X=x] = \binom{n}{x} p^x (1-p)^{100-x}$ //

$$\mu = E[X] := \boxed{\sum_{x=0}^{n}} x \binom{n}{x} p^x (1-p)^{n-x}$$

$\binom{n}{x} = \frac{n!}{(n-x)!\,x!}$
$$= \sum_{x=1}^{n} x \binom{n}{x} p^x (1-p)^{n-x}$$

$n \frac{(n-1)!}{(x-1)!}$
$$= \sum_{x=1}^{n} n \binom{n-1}{x-1} p^x (1-p)^{n-x}$$

$$= n \sum_{x=1}^{n} \binom{n-1}{x-1} p^x (1-p)^{n-x}$$

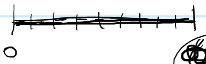
$y = x-1$
$$= n \sum_{y=0}^{n-1} \binom{n-1}{y} p^{y+1} (1-p)^{n-y-1}$$

$$= np \underbrace{\sum_{y=0}^{n-1} \binom{n-1}{y} p^y (1-p)^{n-1-y}}_{= 1}$$

$= np$

$E[X^2] = ?$ $\Rightarrow$ $Var(X) = E[X^2] - E[X]^2$

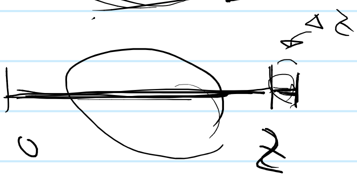
③ Poisson random variable P237



λ : the mean number of "call" per unit interval (time)

Δz λ

$$\lim_{\Delta z \to 0} \frac{g(\Delta z)}{\Delta z} = 0$$



$\lambda \Delta z$

$P_x(z + \Delta z)$

$$P(E_1) = P_x(z)(1 - \lambda \Delta z - o(\Delta z))$$

$$P(E_3) = o(\Delta z)$$

$$= \sum_{i=2,\cdots,x} P_{x-i}(z)\, o(\Delta z)$$

$$\lim_{\Delta z \to 0} \frac{\sum_{i=2,\cdots,x} P_{x-i}(z)\, o(\Delta z)}{\Delta z} = 0$$

$$P_x(z) = P(E_1 \cup E_2 \cup E_3) = P(E_1) + P(E_2) + P(E_3)$$

$$P_0(z) = C e^{-\lambda z}, \quad P_0(0) = 1$$

$$\frac{dP_0(z)}{dz} = -\lambda C e^{-\lambda z} = -\lambda P_0(z)$$

$$P_0(z) = C e^{-\lambda z}$$

$$P_0(z) = Ce$$

$$1 = C$$

$$\frac{dP_x(z)}{dz} + \lambda P_x(z) = \lambda P_{x-1}(z)$$

we want to find out an integrating factor $u(z)$ such that

$$\frac{d}{dz}\left(u(z) P_x(z)\right) = u(z)\frac{dP_x(z)}{dz} + \lambda u(z) P_x(z)$$

It is easy to see $u(z) = e^{\lambda z}$ integrating factor

If $x=1$, then

$$\frac{d\left(e^{\lambda z} P_1(z)\right)}{dz} = \lambda e^{\lambda z} P_0(z)$$

$$= \lambda e^{\lambda z} e^{-\lambda z}$$

$$= \lambda$$

So, $e^{\lambda z} P_1(z) = \lambda z + C$

$$P_1(0) = 0 \Rightarrow C = 0$$

$$e^{\lambda z} P_1(z) = \lambda z$$

$$P_1(z) = \lambda z e^{-\lambda z}$$

Now, we proceed by induction

Claim: $P_x(z) = \dfrac{(\eta z)^x e^{-\eta z}}{x!}$

$x = 0, 1$ it's true

Suppose it holds when $x = n$

$$\frac{d}{dz}\left[e^{\eta z} P_{n+1}(z)\right] = e^{\eta z} \eta P_n(z)$$

$$= \eta \frac{(\eta z)^n}{n!}$$

$$e^{\eta z} P_{n+1}(z) = \frac{(\eta z)^{n+1}}{(n+1)!} + C$$

$P_{n+1}(0) = 0$

$$\Rightarrow P_{n+1}(z) = \frac{(\eta z)^{n+1}}{(n+1)!} e^{-\eta z}$$

$$\mu = E[X] := \sum_{x=0}^{\infty} x \cdot \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= \sum_{x=1}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= \lambda e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!}$$

$$= \lambda e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^y}{y!}$$

$$= \lambda e^{-\lambda} e^{\lambda}$$

$$= \lambda$$

$$\sigma^2 = E[X^2] - E[X]^2 = \lambda$$

Chapter 9. Ideal Models of (Absolutely) Continuous

Random variables

① The Exponential Random variable

What is the distribution of inter_arrival of poisson occurrence?

Putting it another way

We wish to consider the random variable, X, representing the total interval size until we observe the first occurrence of Poisson events since the last observation.

We call this random variable X the exponential random variable.

Consider a random variable $Y(t)$ representing the total numbers of occurrences in the interval $(0,t)$ and $Y(t)$ is a Poisson random variable with $\lambda = 2t$.

Let T be the time to the first occurrence. What is P_T?

first occurrence. What is P_T?

$$F_T(t) = P_T[T \le t] = 1 - P_T[T > t]$$

$$= 1 - P_{Y(t)}[Y(t) = 0]$$

$$= 1 - e^{-\lambda t} \quad t > 0$$

By differentiating it,

$$f_T(t) = \begin{cases} \lambda e^{-\lambda t} & t > 0 \\ 0 & \text{otherwise} \end{cases}$$

mean $\mu = E[X] := \int_{-\infty}^{\infty} t f_T(t)\, dt$

$$= \int_0^{\infty} t \lambda e^{-\lambda t}\, dt$$

$$= -t e^{-\lambda t} \Big|_0^{\infty}$$

$$- \int_0^{\infty} - e^{-\lambda t}\, dt$$

$$= \left[-\frac{1}{\lambda} e^{-\lambda t} \right]_0^{\infty}$$

$$= \frac{1}{\lambda}$$

$$\sigma_X^2 := E[X^2] - (E[X])^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

$$E[X^2] = \int_0^{\infty} t^2 \lambda e^{-\lambda t}\, dt$$

$$= 2\int_0^\infty t e^{-\lambda t}\, dt$$

$$= 2\frac{1}{\lambda}\boxed{\int_0^\infty t \lambda e^{-\lambda t}\, dt}$$

$$= \frac{2}{\lambda^2}.$$

memoryless Property.

$$P_T[T > s+t \mid T > s] = P[T > t]$$

$$= \frac{P_T[\{T > s+t\}, \{T > s\}]}{P_T[T > s]} \qquad T > s+t$$

$$= \frac{P_T[T > s+t]}{P_T[T > s]}$$

$$= \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t}$$

$$= P_T[T > t]$$

Note exponential distribution is the only one satisfying memoryless property.

② Gaussian (Normal) Distribution

The Pdf for a Gaussian random

variable X is

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}, \quad -\infty < x < \infty$$

$$X \sim N(\mu, \sigma^2)$$

Fact (i) $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$

(ii) $E[Z] = 0$

(iii) $E[X] = \mu$

(iv) $Var(Z) = 1$

(v) $Var(X) = \sigma^2$

proof) ① $P[Z \le z] = P\left[\frac{X-\mu}{\sigma} \le z\right]$

$$= P[X \le z\sigma + \mu]$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{z\sigma+\mu} e^{-(x-\mu)^2/2\sigma^2} dx$$

$t = \frac{x-\mu}{\sigma}$

$\sigma dt = dx$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{t^2}{2}} dt$$

(ii) $E[Z] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z e^{-z^2/2} dz$

$$= \frac{1}{\sqrt{2\pi}} \left[-e^{-\frac{z^2}{2}}\right]_{-\infty}^{\infty}$$

$$= 0$$

(iii) $E[X] = E[\mu + \sigma Z]$

$\mu \int_{-\infty}^{\infty} \mu f_T(t)\,dt = E[\mu] + \sigma E[Z]$

$= \mu$

(iv) $Var(Z) = 1$ and

$Var(X) = \sigma^2$

Finally, $\boxed{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2}\, dz \overset{?}{=} 1}$

Claim: $\int_0^{\infty} e^{-z^2/2}\, dz$

$= \frac{\sqrt{2\pi}}{2} = \sqrt{\frac{\pi}{2}}$

Note ~~that~~ the function $e^{-z^2/2}$ does not have an antiderivative that can be written explicitly in terms of elementary functions, so we cannot perform the integration directly

$\left(\int_0^{\infty} e^{-z^2/2}\, dz\right)^2$

$= \left(\int_0^{\infty} e^{-z^2/2}\, dz\right)\left(\int_0^{\infty} e^{-t^2/2}\, dt\right)$

$= \int_0^{\infty}\int_0^{\infty} e^{-t^2/2}\, dt\, e^{-z^2/2}\, dz$

$$= \int_0^\infty \int_0^\infty e^{-t^2/2} e^{-z^2/2} \, dt \, dz$$

$$= \int_0^\infty \int_0^\infty e^{-\frac{(t^2+z^2)}{2}} \, dt \, dz$$

$z = r\cos\theta$
$t = r\sin\theta$

$0 < r < \infty$
$0 < \theta < \frac{\pi}{2}$

$$= \int_0^\infty \int_0^{\frac{\pi}{2}} e^{-\frac{r^2}{2}} \, r \, d\theta \, dr$$

$$= \frac{\pi}{2} \int_0^\infty r e^{-\frac{r^2}{2}} dr$$

$$= \frac{\pi}{2} \left[-e^{-\frac{r^2}{2}} \right]_0^\infty$$

$$= \frac{\pi}{2}$$

Chapter 5 Multidimensional Random Variables.

We are going to investigate situations in which we deal simultaneously with several random variables on the same sample space.

ex) Male students at Yonsei University

X : Height Y : Weight

What is $P[X > 170, Y < 70]$

what is $P[X > 170, Y < 70]$

$$\| $$

$P[\{X > 170\} \cap \{Y < 70\}]$

In general, $P_{XY}[(X,Y) \in A] = ?$

$A \in \mathcal{B}_2$

We are going to cover only

2-d random vector.

The goal is to find the

probability P_{XY} on $\mathcal{B}_2$

(a σ-algebra consisting

of subsets of $\mathbb{R}^2$)

We go through finite or countable

valued random vector with

an example.

(ex) 5.1. tossing a coin 2 times

$\Omega = \{HH, HT, TH, TT\}, \mathcal{A}, P$

Define a random vector

$X = (X_1, X_2)$

by X_1: the number of

heads obtained

in the first toss

X_2: ________

in the second toss

Thus, $X(HH) = (X_1(HH), X_2(HH))$

$= (1, 1)$

$\Omega \xrightarrow{X} \mathbb{R}^2$

$X(HT) = (1, 0)$

$X(TH) = (0, 1)$

$X(TT) = (0, 0)$

So,

$P_X((X \in A))$

$= P(X^{-1}(A))$

$= 0$

$P_X(X_1 = 1, X_2 = 1) = \frac{1}{4}$

$P_X(X_1 = 0, X_2 = 1) = \frac{1}{4}$

$P_X(X_1 = 0, X_2 = 1) = \frac{1}{4}$

$P_X(X_1 = 0, X_2 = 0) = \frac{1}{4}$

What is the probability

$P_X(X_1 > 2, X_2 < -1)$

$= 0$

From P_X, we have Joint pmf

$$f_X(x_1, x_2) = \begin{cases} \frac{1}{4} & x_1 = 1, x_2 = 1 \\ \frac{1}{4} & x_1 = 1, x_2 = 0 \\ \frac{1}{4} & x_1 = 0, x_2 = 0 \\ 0 & \text{otherwise.} \end{cases}$$

$$P[X_1 \in A, X_2 \in B] = \sum_{(x_1, x_2) \in A \times B} f_X(x_1, x_2)$$

$$P[X \in A] = \sum_{x \in A} f_X(x)$$

Naturally, f_X must satisfy

Naturally, f_X must satisfy

① $f_X(x_1, x_2) \geq 0, \ \forall (x_1, x_2) \in \mathbb{R}^2$

② $\sum_{x \in \mathbb{R}^2} f(x) = 1$

The story here goes the same as the case of 1-d random variable.

Thus, we have a natural extension of probability distribution function of random variable for the random vector:

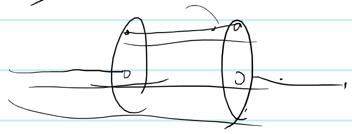
$$F_X(x_1, x_2) = P_X(X_1 \leq x_1, X_2 \leq x_2), \ \forall (x_1, x_2) \in \mathbb{R}^2$$

For some random vector X, it is possible that

$$\forall A \in \mathcal{B}_2, \ P_X(A) = \iint_A f_X(x, y)\,dx\,dy$$

$F(x) = \int_{-\infty}^{x} f(t)\,dt$

Joint pdf.



$$\frac{\partial F_X(x, y)}{\partial x \partial y} = f_X(x, y)$$

f_X must satisfy:

① $f(x_1, x_2) \geq 0$

② $\iint_{\mathbb{R}^2} f(x_1, x_2)\, dx_1 dx_2 = 1$

Now, at this point, it is natural to ask if we get P_{X_1} and P_{X_2} from P_X.

Idea: Reconsider the example tossing coin twice

$$\Omega = \{HH, HT, TH, TT\}$$

$$\begin{aligned} P_{X_1}[X_1 = 1] &= P[\{HH, HT\}] \\ &= P[\{HH\} \cup \{HT\}] \\ &= P[\{HH\}] + P[\{HT\}] \\ &= P_X(X_1=1, X_2=1) + P_X(X_1=1, X_2=0) \\ &= f_X(1,1) + f_X(1,0) \end{aligned}$$

So, $f_{X_1}(1) = \sum_{x_2} f_X(1, x_2) = \frac{1}{2}$

$f_{X_1}(0) = \sum_{x_2} f_X(0, x_2) = \frac{1}{2}$

In conclusion, for a random vector $X = (X_1, X_2)$, the probability distribution of X_1 and X_2

vector $X = (X_1, X_2)$, the probability distribution of X_1 and X_2

are

$$f_{X_1}(x_1) = \begin{cases} \sum_{x_2} f_X(x_1, x_2) \\ 0 \quad \text{o.w.} \end{cases}$$

$$f_{X_2}(x_2) = \sum_{x_1} f_X(x_1, x_2)$$

, respectively

We call f_{X_1}, f_{X_2} marginal pmf of X. (absolutely)

For the case of continuous random vector,

marginal pdf

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f_X(x_1, x_2)\, dx_2$$

So, $P(a_1 < X_1 < a_2)$

$$= \int_{a_1}^{a_2} \int_{-\infty}^{\infty} f_X(x_1, x_2)\, dx_2\, dx_1$$

ex

Independent Random variables

Consider tossing a coin two times

$$\Omega = \{HH, HT, TH, TT\}$$

$A = \{HH, HT\}$ $B = \{HH, TH\}$

then, $P(A \cap B) = P(\{HH\})$

$$= \frac{1}{4}$$

$$P(A) = \frac{1}{2}$$

$$P(B) = \frac{1}{2}$$

$$p(B) = \frac{1}{2}$$

$X_1 = \#$ of H in the first

$X_2 = \#$ of H in the second

$$P_{(X_1, X_2)}(\underbrace{X_1 \in \{1\}}_{X_1 = 1}, \underbrace{X_2 \in \{1\}}_{X_2 = 1})$$

$$= P[\{HH, HT\} \cap \{HH, TH\}]$$

$$= P[\{HH\}] = \frac{1}{4}$$

$$P_{X_1}(X_1 \in \{1\}) = P[\{HH, HT\}]$$

$$= \frac{1}{2}$$

$$P_{X_2}(X_2 \in \{1\}) = P[\{HH, TH\}]$$

$$= \frac{1}{2}$$

Definition Random variables $X_1, \cdots, X_n$ defined on the same probability space $(\Omega, \mathcal{A}, P)$ are said to be mutually independent

if $P(X_1 \in B_1, X_2 \in B_2, \cdots, X_n \in B_n)$

$= P(X_1 \in B_1) \cdots P(X_n \in B_n)$

$B_1, B_2, \cdots, B_n \in \mathcal{B}$

theorem Suppose that $X_1, \cdots, X_n$ are random variables, all defined on the same probability space $(\Omega, \mathcal{A}, P)$. Then, $X_1, \cdots, X_n$ are mutually independent iff for all

•f u r l t i o u X l ^ , 쨋 (l 0 , 蜜 k m e d a 빠 V 억 s k 페이지 26

$x_1, \dots, x_n \in \mathbb{R}$,

$F_X(x_1, x_2, \dots, x_n) = F_{X_1}(x_1) \cdots F_{X_n}(x_n)$

Corollary Let (X, Y) be a random vector with joint pdf or pmf $f_X(x)$ and $f_Y(y)$. Then X and Y are mutually independent ~~$f(x,y) = f_X(x) f_Y(y)$~~ iff for $\forall x, y \in \mathbb{R}$

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

Corollary Let X, Y be ~~random~~ independent random variables. Then

$$E[g(X) h(Y)] = E[g(X)] E[h(Y)]$$

① $E[g(X,Y)] := \int_\Omega g(X(\omega), Y(\omega)) dP(\omega)$

$= \iint_{\mathbb{R}^2} g(x,y) \, dP(x,y)$

$= \iint_{\mathbb{R}^2} g(x,y) f(x,y) \, dx \, dy$

or

$= \sum_{\mathbb{R}^2} g(x,y) f(x,y)$

② Covariance and Correlation

Independence : Probabilistic relationship between random variables

Then, natural questions are are ① what the relation is

② How strong or weake the relation is if X, Y are dependent.

ex Male students at Yonsei University

X: Height Y: weight

In this example, large values of X tend to be observed with large values of Y and small values of X with small values of Y. If $X > \mu_X$, then $Y > \mu_Y$ is likely to be true and the product $(X - \mu_X)(Y - \mu_Y)$ will be positive

If $X < \mu_X$, then $Y < \mu_Y$ is likely to be true and the product $(X - \mu_X)(Y - \mu_Y)$ will again be positive.

Thus, the sign of $(X - \mu_X)(Y - \mu_Y)$ gives us information regarding the relationship between X and Y

Tendency: The covariance of X and Y is the number defined by

$$Cov(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

How strong or weak is the relation

Tendency: The correlation of X and Y is the numbers defined by

$$\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

(X,Y): independent, g, h.

$$E[g(X)\, h(Y)] = E[g(X)]\, E[h(Y)]$$

$$\begin{aligned}
E[g(X)\, h(Y)] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)\, f(x,y)\, dx\, dy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)\, f(x)f(y)\, dx\, dy \\
&= \int_{-\infty}^{\infty} h(y)f(y) \underbrace{\left(\int_{-\infty}^{\infty} g(x)f(x)\, dx\right)}_{= E[g(X)]} dy \\
&= E[g(X)] \underline{\int_{-\infty}^{\infty} h(y)f(y)\, dy} \\
&= E[g(X)]\, E[h(Y)]
\end{aligned}$$