[Probability Theory]

Random Variable X

Discrete X,   Continuous Y

Probability mass function f_X Probability distribution function f_Y

p.m.f   p.d.f

$$f_X : \underbrace{\Omega}_{\text{Sample Space}} \to \mathbb{R} \geq 0$$

$$f_Y : \overset{/\!/\mathbb{R}}{\Omega} \to \mathbb{R} \geq 0$$

$$\sum_{x \in \Omega} f_X(x) = 1$$

$$\int_\Omega f_Y(Y)dy = 1$$

Ex.)

$$X = \begin{cases} 1 & (\text{prob } 1/3) \\ -1 & (\text{prob } 1/3) \\ 0 & (\text{prob } 1/3) \end{cases}$$

$$f_X(1) = f_X(-1) = f_X(0) = \frac{1}{3}$$

Ex.) $f_Y(y) = 1$ for all $y \in \{0, 1\}$

probability distribution function of random variable $[0, 1]$

Probability of event

$$P(A) = \sum_{x \in A} f_X(x)$$
$$= \int_A f_Y(y)\,dy$$

Expectation

$$\mathbb{E}[X] = \sum_{x \in \Omega} x f_X(x) \qquad \mathbb{E}[Y] = \int_\Omega y f_Y(y)\,dy$$

Two random variables X_1, X_2 are independent if

$P(X_1 \in A \text{ and } X_2 \in B) = P(X_1 \in A)\,P(X_2 \in B)$ for all events A and B

<independence>

- $\boxed{\text{Mutually independent events}}$
- Pairwise independent events

ex) $\overbrace{X_1, X_2, X_3}^{\text{indep.}}$

indep. indep.

independent X

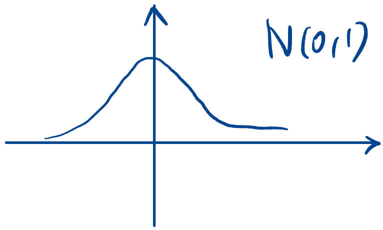
Normal Distribution

probability density function

Continuous random variable has normal distribution $N(\mu, \sigma)$

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

μ: mean $\quad \sigma$: variance



P: Price

Ex) $P_n - P_{n-1} \sim N(0,1)$

Not a good model

Instead we want the "relative difference" to be normally distributed.

$$\frac{P_n - P_{n-1}}{P_n} \sim N(0,1)$$

(we want the percentage change to be normally distributed.)

Q. What does the distribution of P_n look like?

Lognormal Distribution Y.

Log-normal random variable Y such that $\log Y$ is randomly distributed

Theorem Change of Variable

Suppose X, Y are random variables such that

$$P(X \le x) = P(Y \le h(x)) \text{ for all } x$$

Then, $f_X(x) = f_Y(h(x)) \cdot h'(x)$

X: log-normal distribution parameters μ, σ (mean X, variance X)

Y: normal distribution mean: μ var: σ

$$P(X \le x) = P(Y \le \log x)$$

$$f_X(x) = f_Y(\log x)\frac{1}{x} = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\log x - \mu)^2}{2\sigma^2}} \quad (x > 0)$$

probability mass function

Log-normal random variable Y such that Y is normally distributed.

Poisson distribution

Exponential distribution

Exponential family of distribution

A distribution belongs to exponential family

if there exists vector θ that parameterizes distribution

such that $f^{(\theta)}(x) = h(x)c(\theta)\exp\left(\sum_{i=1}^{k} w_i(\theta)t_i(x)\right)$

$h(x)$, $t_i(x)$ depends only on X

$c(\theta)$, $w_i(\theta)$ depends only on θ

$$\frac{1}{x\sigma\sqrt{2\pi}} e^{\frac{(\log x - \mu)^2}{2\sigma^2}} = \frac{1}{x} \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{(\log x)^2}{2\sigma^2} + \frac{\mu \log x}{\sigma^2} - \frac{\mu^2}{2\sigma^2}}$$

$\theta = (\mu, \sigma)$

$h(x) = \frac{1}{x}$ $\quad c(\theta) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{\mu^2}{2\sigma^2}}$

$t_1(x) = (\log x)^2$

$w_1(x) = -\frac{1}{2\sigma^2}$

$t_2(x) = \log x$

$w_2(x) = \frac{\mu}{\sigma^2}$

Given a random variable

1. Study 'Statistics'

   k-th moment of random variable $\mathbb{E}X^k$

   Moment generating function

2. Study long-term (large-scale) behavior

   - Law of large numbers
   - Central limit theorem

Moment Generating function

has all the information about random variables.

$$M_X(t) = \mathbb{E}\, e^{tX} \quad t \in \mathbb{R}$$

Remark

- Does not necessarily exist.
- Log-normal distribution does not have moment generating function

$$\frac{d^{(k)} M_X}{dt^{(k)}}(t) = \mathbb{E}X^k \quad \text{for all } k \in \mathbb{N}$$

$$M_X(0) = \sum_{k=0}^{\infty} \frac{t^k}{k!} m_k \quad m_k = \mathbb{E}X^k$$

THM (i) If X, Y have the same moment generating function, then X, Y have the same distribution.

Remark: It does not imply that all random variables with identical k-th moments ($k \in \mathbb{N}$) have the same distribution.

THM

$X_1, X_2, \cdots$ is a sequence of random variables such that

$$M_{X_i}(t) \to M_X(t)$$

for some random var X for all t

For all x,

$$P(X_i \leq x) \longrightarrow P(X \leq x)$$

Law of large numbers

THM (Weak law of large numbers)

Let $X_1, X_2, \cdots, X_n$ be independent random variables with identical distributions (i, i, d)

mean : μ , variance : σ^2

$$X = \frac{1}{n}(X_1, X_2, \cdots, X_n)$$

average

Then, $\forall \varepsilon > 0$, $P(|X - \mu| \geq \varepsilon) \to 0$ as $n \to \infty$

Ex $\boxed{\text{Casino}}$

Playing Blackjack under optimal strategy 48% chance of winning.

From the casino's point of view, they are taking enormous n.

Proof

$$\mathbb{E}[\bar{x}] = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right]$$

$$= \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}X_i = \mu$$

$$\mathbb{V}[\bar{x}] = \mathbb{E}[(\bar{X}-\mu)^2]$$

$$= \mathbb{E}\left[\left(\frac{1}{n}\Sigma X_i - \mu\right)^2\right]$$

$$= \mathbb{E}\left[\frac{1}{n^2}\sum_{i=1}^{n}(X_i-\mu)^2\right]$$

$$= \frac{\sigma^2}{n}$$ affects your variance by $\frac{1}{n}$

$$\varepsilon^2 P(|\bar{X}-\mu| \geq \varepsilon) \leq \mathbb{V}[\bar{X}] = \frac{\sigma^2}{n}$$

$$P(|\bar{X}-\mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

Want to be sure that $|\bar{x}-\mu| < 0.1$