Advanced Mathematical Methods

WS 2023/24

4 Mathematical Statistics

Dr. Julie Schnaitmann

Department of Statistics, Econometrics and Empirical Economics

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Wirtschafts- und Sozialwissenschaftliche Fakultät

Outline: Mathematical Statistics

Readings

- A. Papoulis and A. U. Pillai. *Probability, Random Variables and Stochastic Processes*.
  Mc Graw Hill, fourth edition, 2002, Chapters 1-4

Online References

MIT Course on Probabilistic Systems Analysis and Applied Probability (by John Tsitsiklis)

- Discrete RVs I: Concept of random variables, probability mass function, expected value, variance
  https://www.youtube.com/watch?v=3MOahpLxj6A
- Continuous RVs: probability density function, cumulative distribution function, expected value, variance
  https://www.youtube.com/watch?v=mHfn_7ym6to
- Discrete RVs II: Functions of RV, conditional probabilities, specific distribution, total expectation theorem, joint probabilities
  https://www.youtube.com/watch?v=-qCEoqpwjf4

4.1 Random Variables

A random variable X takes on real numbers according to some distribution.

There are two types of random variables:

1. discrete random variables
   - e.g. coin toss, number of baskets scored out of n trials
2. continuous random variables
   - e.g. financial returns

1. Bernoulli

X success $= 1$ T $\Pr(X=1)=\pi$

failure $= 0$ H $\Pr(X=0) = 1-\Pr(X=1)$

$= 1-\pi$

4.1 Random Variables

A random variable X takes on real numbers according to some distribution.

There are two types of random variables:

countably many outcomes

1. discrete random variables
    - e.g. coin toss, number of baskets scored out of n trials
2. continuous random variables
    - e.g. financial returns

infinitely many outcomes

4.1 Random Variables

Discrete random variables

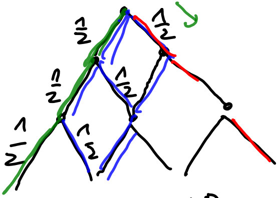
(2) Binomial distribution: n independent Bernoulli trials

X counts how many times E was successful out of n trials

H / \ T

tails

ex: $n=3$



$X=0$ 1 $X=0$

$X=1$ 3 $X=1$

$X=2$ 3

$X=3$ 1

$$\Pr(X=k) = \pi^k \cdot (1-\pi)^{(n-k)} \binom{n}{k}$$

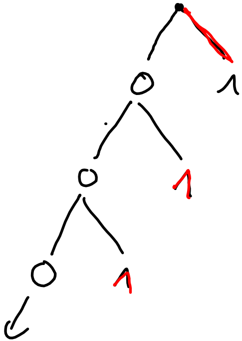
4.1 Random Variables

Discrete random variables

3 Geometric distribution

"How many times do I fail before I succeed?"

$$Pr(X = k) = (1-\pi)^k \cdot \pi$$



$X = 0$

$X = 1$

$X = 2$

$\vdots$

4.1 Random Variables

Discrete random variables

4.1 Random Variables

Random sample

n random variables

$\{X_1, X_2, \ldots, X_n\}$ is called a random sample if

1. all draws X_i are **independent**
2. and drawn from the same distribution, i.e. they are **identically distributed**

$\Rightarrow$ the draws are **independently and identically distributed** in short **iid**

ex: coin toss: $n=10$ $\{X_1, X_2, \ldots, X_{10}\}$

realization: $\{x_1, x_2, \ldots, x_{10}\} = \{T, H, H, T, \ldots, H\}$

4.2 Cumulative Distribution Functions

Probability distribution function: discrete case (pdf)

pdf

realization/ specific outcome

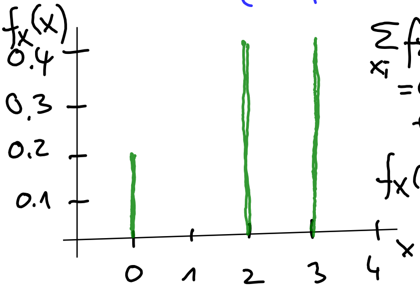
$$f_X(x_i) = P(X = x_i) = Pr(X = x_i)$$

r.v.

ex: $f_X(x_i) = \begin{cases} 0.2 & x=0 \\ 0.4 & x=2 \\ 0.4 & x=3 \end{cases}$

requirements:

- $0 \le P(X = x_i) \le 1$
- $\sum_{x_i} f_X(x_i) = 1$



$\sum_{x_i} f_X(x_i) = 0.2 + 0.4 + 0.4 = 1$

$f_X(1) = 0$

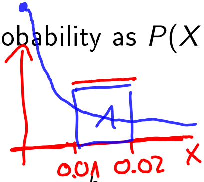
4.2 Cumulative Distribution Functions

(Probability) Density function: continuous case

$f_X(x)$ is not a probability as $P(X = x) = 0$

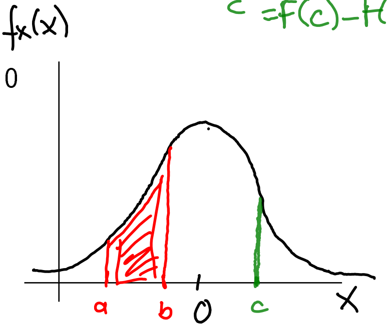
not a probability = point mass is zero

$P(X=c) = \int_c^c f_X(x)\,dx = F(c) - F(c) = 0$



requirements:

- $P(a \leq X \leq b) = \int_a^b f_X(x)\,\mathrm{d}x \geq 0$
- $\int_{-\infty}^{\infty} f_X(x)\,\mathrm{d}x = 1$
- $f_X(x) \geq 0$ non-negative



4.2 Cumulative Distribution Functions

Definition: Cumulative distribution function

The cumulative distribution function (cdf) of a random variable X is defined to be the function $F_X(x) = P(X \leq x)$, for $x \in \mathbb{R}$.

discrete:

$$F_X(x_i) = \sum_{X \leq x_i} f_X(x_i) = P(X \leq x_i)$$

continuous:

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,\mathrm{d}t = P(X \leq x)$$

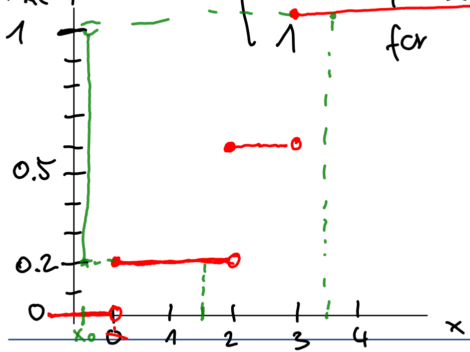
4.2 Cumulative Distribution Functions

Discrete random variable

ex: $f_X(x_i) = \begin{cases} 0.2 & X=0 \\ 0.4 & x=2 \\ 0.4 & X=3 \end{cases}$

$F_X(x_i) = \sum_{X \le x_i} f_X(x_i)$

$F_X(x_i) = \begin{cases} 0 & \text{for } X<0 \\ 0.2 & \text{for } 0 \le X < 2 \\ 0.6 & \text{for } 2 \le X < 3 \\ 1 & \text{for } X \ge 3 \end{cases}$



$Pr(1.5 \le X \le 3.5)$

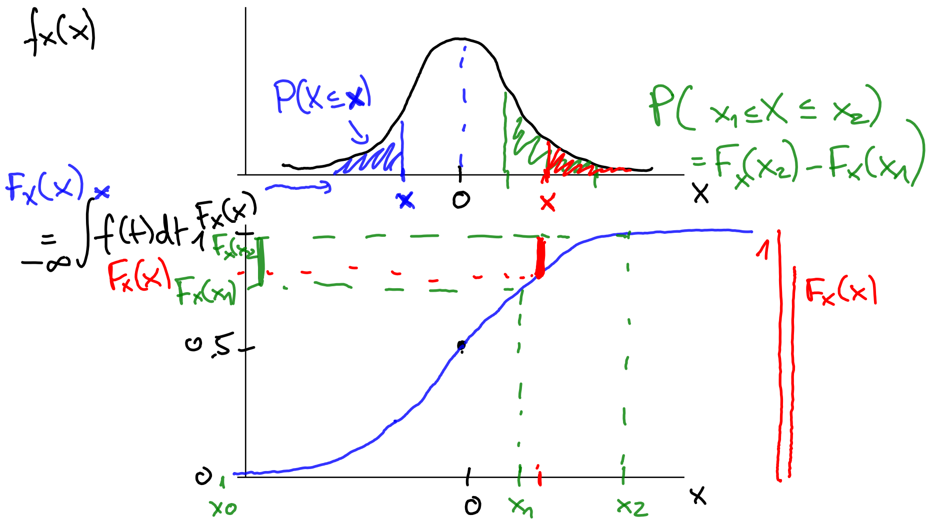
$= Pr(X \le 3.5) - Pr(X \le 1.5)$

$= F_X(3.5) - F_X(1.5)$

$= 0.8$

4.2 Cumulative Distribution Functions

Continuous random variable



4.2 Cumulative Distribution Functions

Properties

1) $F_X(+\infty) = 1$; $F_X(-\infty) = 0$

2) $F_X(x)$ is a nondecreasing function of x:
   if $x_1 < x_2$, $F_X(x_1) \leq F_X(x_2)$
   note: the event $\{X \leq x_1\}$ is a subset of $\{X \leq x_2\}$

3) if $F_X(x_0) = 0$, then $F_X(x) = 0 \quad \forall \quad x \leq x_0$

4.2 Cumulative Distribution Functions

Properties

4) $P(X > x) = 1 - F_X(x)$
events $\{X \leq x\}$ and $\{X > x\}$ are mutually exclusive and
$\{X \leq x\} \cup \{X > x\} = \Omega$

5) $F_X(x)$ is continuous from the right:
$\lim_{x \to a^+} F_X(x) = F_X(a)$

6) $P(x_1 \leq X \leq x_2) = F_X(x_2) - F_X(x_1) = \int_{x_1}^{x_2} f_X(x)\, dx$

4.3 Expectation, Variance and Moments

Expectations of a random variable

weighted average $\frac{\cancel{E}(X)}{\cancel{E}(Y)}$

$$E[X] = \begin{cases} \sum\limits_{x_i} x_i f_X(x_i) & \text{if } x \text{ is discrete} \\ \int\limits_{-\infty}^{\infty} x f_X(x)\,\mathrm{d}x & \text{if } x \text{ is continuous} \end{cases}$$

ex: coin toss: X T: $X=1$, $\pi = \frac{1}{2}$ H: $X=0$, $1-\pi = \frac{1}{2}$

If $g(X)$ a measurable function of x, then:

$$E[g(X)] = \begin{cases} \sum\limits_{x_i} g(x_i) f_X(x_i) & \text{if } x \text{ is discrete} \\ \int\limits_{-\infty}^{\infty} g(x) f_X(x)\,\mathrm{d}x & \text{if } x \text{ is continuous} \end{cases}$$

$$E(X) = 1 \cdot P(X=1) + 0 \cdot P(X=0)$$
$$= 1 \cdot \pi + 0 \cdot (1-\pi) = \pi = \frac{1}{2}$$

4.3 Expectation, Variance and Moments

Expectations of a random variable

$$E[X] = \begin{cases} \sum_{x_i} x_i f_X(x_i) & \text{if } x \text{ is discrete} \\ \int_{-\infty}^{\infty} x f_X(x) \mathrm{d}x & \text{if } x \text{ is continuous} \end{cases}$$

If $g(X)$ a measurable function of x, then:

$$E[g(X)] = \begin{cases} \sum_{x_i} g(x_i) f_X(x_i) & \text{if } x \text{ is discrete} \\ \int_{-\infty}^{\infty} g(x) f_X(x) \mathrm{d}x & \text{if } x \text{ is continuous} \end{cases}$$

ex: $g(x) = x^2$

coin toss

$x_1^2 = 1^2 = 1$ $\quad x_2^2 = 0^2 = 0$

$E(X^2) = 1^2 \cdot P(X=1) + 0^2 \cdot P(X=0)$

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

$= \pi$

4.3 Expectation, Variance and Moments

Calculation rules

constant not a r.v.

- $E[a] = a$ $= a \cdot \underbrace{P(X=a)}_{=1} = a$

$$E(a) = \int_{-\infty}^{\infty} a \, f_X(x) \, dx = a \underbrace{\int_{-\infty}^{\infty} f_X(x) \, dx}_{=1} = a$$

- $E[bX] = b \cdot E[X]$
- linear transformation: $E[a + bX] = a + bE[X]$
- $E[g_1(X) + g_2(X)] = E[g_1(X)] + E[g_2(X)]$

4.3 Expectation, Variance and Moments

Calculation rules

- $E[a] = a$
- $E[bX] = b \cdot E[X]$ $\quad = \int_{-\infty}^{\infty} b \cdot x \, f_X(x)\, dx = b \underbrace{\int_{-\infty}^{\infty} x\, f_X(x)\, dx}_{=E(X)}$

  $= b \cdot E(X)$
- linear transformation: $E[a + bX] = a + bE[X]$
- $E[g_1(X) + g_2(X)] = E[g_1(X)] + E[g_2(X)]$

4.3 Expectation, Variance and Moments

Calculation rules

- $E[a] = a$
- $E[bX] = b \cdot E[X]$
- linear transformation: $E[a + bX] = a + bE[X]$

seperate integrals

check if linear

$$= \int_{-\infty}^{\infty} \underbrace{(a+bx)}_{g(x)} \cdot f_X(x)\,dx = \int_{-\infty}^{\infty} a \cdot f_X(x)\,dx + \int_{-\infty}^{\infty} b x\, f_X(x)\,dx$$

- $E[g_1(X) + g_2(X)] = E[g_1(X)] + E[g_2(X)]$

look for familiar terms

$$= a \int_{-\infty}^{\infty} \underbrace{f_X(x)\,dx}_{=1} + b \int_{-\infty}^{\infty} \underbrace{x f_X(x)\,dx}_{=E(X)} = a + bE(X)$$

constants in front of integrals

4.3 Expectation, Variance and Moments

Calculation rules

- $E[a] = a$
- $E[bX] = b \cdot E[X]$
- linear transformation: $E[a + bX] = a + bE[X]$
- $E[g_1(X) + g_2(X)] = E[g_1(X)] + E[g_2(X)]$

$$E(\overbrace{bX} + X^2) = E(bX) + E(X^2) = bE(X) + E(X^2)$$

$E(\cdot)$ linear operator

4.3 Expectation, Variance and Moments

Variance of a random variable

$(a+b)^2 = a^2 + 2ab + b^2$

Let $g(X) = (X - E[X])^2$

Variance operator is <u>not</u> linear

$$Var[X] = \sigma^2 = E[(X - E[X])^2] = E(g(x))$$

$$= \begin{cases} \sum\limits_{x_i} (x_i - E[X])^2 f_X(x_i) & \text{if } x \text{ is discrete} \\ \int\limits_{-\infty}^{\infty} (x - E[X])^2 f_X(x) \mathrm{d}x & \text{if } x \text{ is continuous} \end{cases}$$

$Var(a + bX) \neq a + b\,Var(X)$

$= b^2 Var(X)$

4.3 Expectation, Variance and Moments

Calculation rules

- $Var[a] = 0$

$E(a) = a$

$Var(a) = E\left[\left(a - \underbrace{E(a)}_{=a}\right)^2\right] = 0$

- $Var[X + a] = Var[X]$
- $Var[bX] = b^2 Var[X]$
- $Var[a + bX] = b^2 Var[X]$

important result:

$Var[X] = E[X^2] - E[X]^2$

4.3 Expectation, Variance and Moments

Calculation rules

- $Var[a] = 0$
- $Var[X + a] = Var[X]$

$$E\left[\left((X+a) - \underbrace{E(X+a)}_{E(X)+a}\right)^2\right] = E\left[\left(X + \cancel{a} - E(X) \cancel{- a}\right)^2\right]$$

$$E(X)+a = E(X) + \underbrace{E(a)}_{=a}$$

$$= E\left[\left(X - E(X)\right)^2\right]$$

$$= Var(X)$$

- $Var[bX] = b^2 Var[X]$
- $Var[a + bX] = b^2 Var[X]$

important result:

$Var[X] = E[X^2] - E[X]^2$

4.3 Expectation, Variance and Moments

Calculation rules

- $Var[a] = 0$
- $Var[X + a] = Var[X]$
- $Var[bX] = b^2 Var[X]$
- $Var[a + bX] = b^2 Var[X]$

important result:

$Var[X] = E[X^2] - E[X]^2$

4.3 Expectation, Variance and Moments

Calculation rules

- $Var[a] = 0$
- $Var[X + a] = Var[X]$
- $Var[bX] = b^2 Var[X]$
- $Var[a + bX] = b^2 Var[X]$

$$= E\left[\left(a+bX - \underline{E(a+bX)}\right)^2\right] = E\left[\left(\cancel{a}+bX - \cancel{a} - bE(X)\right)^2\right] = E\left[\left(b(X-E(X))\right)^2\right] = b^2 E\left[(X-E(X))^2\right] = b^2 Var(X)$$

$$= E(a) + E(bX)$$
$$= a + bE(X)$$

Important result:

$Var[X] = E[X^2] - E[X]^2$

4.3 Expectation, Variance and Moments

Calculation rules

- $Var[a] = 0$
- $Var[X + a] = Var[X]$
- $Var[bX] = b^2 Var[X]$
- $Var[a + bX] = b^2 Var[X]$

important result:

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

Variance

r.v. constant

$$\mathrm{Var}(X) = E\left[(X - E(X))^2\right]$$

multiply out

$$= \int_{-\infty}^{\infty} (x - E(x))^2 f_X(x)\,dx = \int (x^2 - 2xE(x) + E(x)^2) f_X(x)\,dx$$

seperate integrals

constants outside

$$= \int x^2 f_X(x)\,dx - \int 2x E(x) f_X(x)\,dx + \int E(x)^2 f_X(x)\,dx$$

$$= \int x^2 f_X(x)\,dx - 2E(x)\underbrace{\int x f_X(x)\,dx}_{=E(X)} + E(x)^2 \underbrace{\int f_X(x)\,dx}_{=1}$$

familiar terms

$$= E(X^2) - 2\underbrace{E(x)\cdot E(x)}_{E(X)^2} + E(x)^2 \cdot 1$$

$$= E(X^2) - E(X)^2$$

4.3 Expectation, Variance and Moments

Standardization of a random variable X

$$Y = X - E(X) = X - \mu$$

center $E(Y) = 0$

$Var(Y) = Var(X)$

Let

$$Z = \frac{X - E(X)}{\sqrt{Var(X)}}$$

$$g(X) = \frac{X-\mu}{\sigma} = Z$$

$a + bX$ standardization

$$Z = \frac{X-\mu}{\sigma} = \frac{-\mu}{\sigma} + \frac{1}{\sigma}X = a + bX$$

$\Rightarrow \quad E[Z] = 0 \quad$ and $\quad Var[Z] = 1$

$E(Z) = 0$ $\qquad$ $Var(Z) = 1$

4.3 Expectation, Variance and Moments

Standardization of a random variable X

$$E\left[\frac{X-\mu}{\sigma}\right] = \frac{1}{\sigma} E[X-\mu] = \frac{1}{\sigma}\left[\underbrace{E(X) - \underbrace{E(\mu)}_{=\mu}}_{=0}\right] = 0$$

Let

$$g(X) = \frac{X-\mu}{\sigma} = Z$$

$$Var\left[\frac{X-\mu}{\sigma}\right] = \frac{1}{\sigma^2} Var[X-\mu] = \frac{1}{\sigma^2} \underbrace{Var(X)}_{=\sigma^2} = 1$$

$$Z = \frac{X-\mu}{\sigma} = \frac{-\mu}{\sigma} + \frac{1}{\sigma} X$$

$\Rightarrow \quad E[Z] = 0$ and $Var[Z] = 1$

4.3 Expectation, Variance and Moments

Chebychev Inequality

For any random variable X with finite expected value μ and finite variance $\sigma^2 > 0$ and a positive constant k

$\sqrt{Var(X)} = \sigma$

$$P(\mu - k\sigma \leq X \leq \mu + k\sigma) \geq 1 - \frac{1}{k^2}$$

$E(X) = \mu$

convergence in probability $\xrightarrow{P}$

4.3 Expectation, Variance and Moments

Skewness and Kurtosis

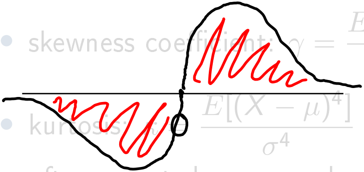
Central moments of a random variable:

$$\mu_r = E[(X-\mu)^r] = \int_{-\infty}^{\infty} (x - E(x))^r \cdot f_x(x)\,dx$$

→ integral diverges

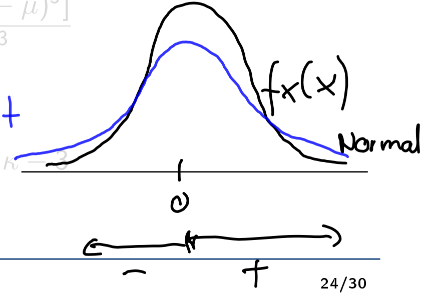
as r grows, μ_r tends to explode

Solution: normalization



- skewness coefficient: $\gamma = \dfrac{E[(X-\mu)^3]}{\sigma^3}$
- kurtosis: $\kappa = \dfrac{E[(X-\mu)^4]}{\sigma^4}$

  often reported as excess kurtosis $\kappa - 3$



$$E(x) = \int_{-\infty}^{\infty} x\, f_x(x)\,dx$$

+

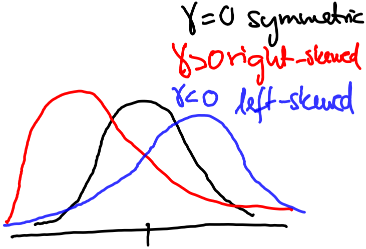
4.3 Expectation, Variance and Moments

Skewness and Kurtosis

Central moments of a random variable:

$$\mu_r = E[(X-\mu)^r]$$

as r grows, μ_r tends to explode

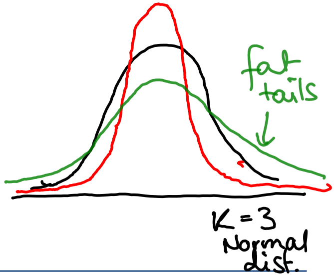


<u>Solution</u>: normalization

- skewness coefficient: $\gamma = \dfrac{E[(X-\mu)^3]}{\sigma^3}$
- kurtosis: $\kappa = \dfrac{E[(X-\mu)^4]}{\sigma^4}$

  often reported as excess kurtosis $\kappa - 3$

+ leptokurtic
− platykurtic



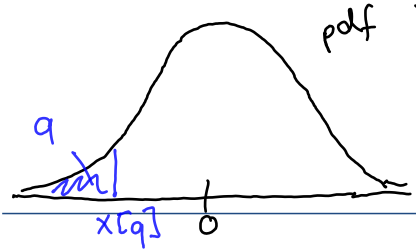
4.4 Quantile

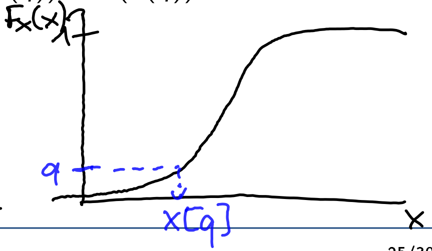
Quantile

q% of the probability mass of a random variable is left of $x(q)$.

Example: Risk measure Value-at-risk (VaR)

$$q = P(X \leq x(q)) = F(x(q))$$





4.5 Specific probability distributions

The normal distribution

X is a Gaussian or normal random variable with parameters μ and σ^2 if its density function is given by

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) = f_X(x; \mu, \sigma)$$

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

Linear transformation is also normally distributed:

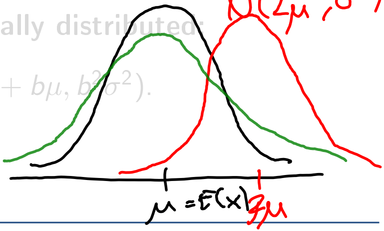
If $X \sim N(\mu, \sigma^2)$, then $a + bX \sim N(a + b\mu, b^2\sigma^2)$.

standardization:

$$Z = \frac{X - E(X)}{\sqrt{Var(X)}} = \frac{X-\mu}{\sigma}$$

$N(0,1)$

$$X = Z \cdot \sigma + \mu$$



4.5 Specific probability distributions

The normal distribution

X is a Gaussian or normal random variable with parameters μ and σ^2 if its density function is given by

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

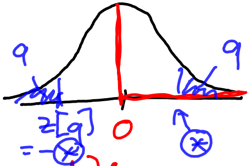
denoted $X \sim N(\mu, \sigma^2)$

Linear transformation is also normally distributed:

If $X \sim N(\mu, \sigma^2)$, then $a + bX \sim N(\underbrace{a + b\mu}_{\substack{E(a+bX) \\ = a+bE(X)}}, \underbrace{b^2\sigma^2}_{\substack{Var(a+bX) \\ = b^2 Var(X)}})$.

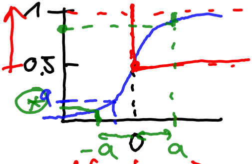
4.5 Specific probability distributions

Standardization of X leads to standard normal distribution:



$$a = -\frac{\mu}{\sigma} \quad , \quad b = \frac{1}{\sigma}$$

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



cdf:

1. standardize

2. table • table

⊛ = 1 − •

$F(-a) = 1 - F(a)$

quantile

1. quantile of z

z[q]

2. retransform x[q] $X = z \cdot \sigma + \mu$

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)$$

Thus, if $X \sim N(\mu, \sigma)$, then $f(x) = \frac{1}{\sigma}\Phi\left(\frac{x-\mu}{\sigma}\right)$.

4.5 Specific probability distributions

Normal distribution

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

symmetry:

$\Phi(-z) = 1 - \Phi(z)$

ex: $P(X \le 2)$

standardization

$$P\Big(\underbrace{\frac{X-\mu}{\sigma}}_{Z} \le \frac{2-\mu}{\sigma}\Big) = P\Big(Z \le \frac{2-\mu}{\sigma}\Big) = \Phi\Big(\frac{2-\mu}{\sigma}\Big)$$

quantiles:

$z[q] = -z[1-q]$

$z[q]$ from table

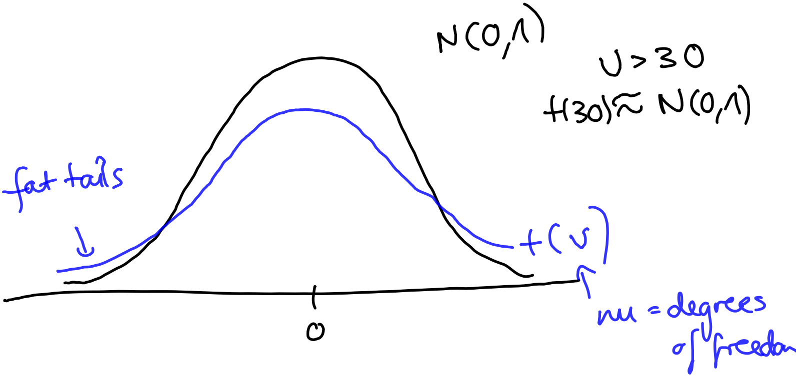
$x[q] = z[q] \cdot \sigma + \mu$

ex:

$z[0.05] = -z[0.95]$

4.5 Specific probability distributions

t-distribution



4.5 Specific probability distributions

The χ^2 distribution:

X is said to be $\chi^2(n)$ with n degrees of freedom if

$$f_X(x) = \begin{cases} \dfrac{x^{\frac{n}{2}-1}}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} e^{-\frac{x}{2}} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

If $Z \sim N(0,1)$, then $x = Z^2 \sim \chi^2(1)$.

If Z_i are **iid** $N(0,1)$, then $\sum_{i=1}^{n} Z_i^2 \sim \chi^2(n)$.

$Z_1, Z_2 \sim N(0,1)$

$Z_1^2 + Z_2^2 \sim \chi^2(2)$