Normal distribution

A cont. RV X is said to have a normal distribution with parameters μ and σ, where $\mu \in \mathbb{R}$ $\sigma \in \mathbb{R}, \sigma > 0$, if the density function of X is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}$$

Properties:

1) Symmetric around mean
$$f(\mu + x) = f(\mu - x)$$

2) Median (= 0,5 quantile) coincides with mean

3) 3 sigma rule

Standard normal distribution

normal distribution with parameters
$\mu = 0 \quad \sigma = 1$

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \qquad z \in \mathbb{R}$$

Theorem 1 if $X \sim N(\mu, \sigma^2)$ and $Y = \frac{X-\mu}{\sigma} \Rightarrow Y \sim N(0,1)$

Proof:

$$P(Y \le x) = P\left(\frac{X-\mu}{\sigma} \le x\right) = P(X \le \mu + \sigma x) =$$

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

$u = \frac{t-\mu}{\sigma}$
$du = \frac{1}{\sigma} dt \Rightarrow dt = \sigma du$
when $t = \mu + \sigma x$
$u = x$

$$= \frac{1}{\cancel{\sigma}\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} \cancel{\sigma} du = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} du = \Phi(x)$$

we proved Y has distribution function $\Phi(x)$, that is, that $Y \sim N(0,1)$

$$EX = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} dx = \int_{-\infty}^{0} + \int_{0}^{\infty} = I_1 + I_2$$

I_2 - exists (how to show it?)

substitute $x = -y$ $\quad dx = -dy$

$y = -x$

when $x = -\infty$

$y = +\infty$

$$I_1 = \int_{0}^{+\infty} (-y) \frac{1}{\sqrt{2\pi}} e^{-\frac{(-y)^2}{2}} (-1) dy = \int_{0}^{+\infty} y \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy$$

$I_1 = -I_2 \implies$ mean is 0

point about odd

$$Var X = E(X^2) - (EX)^2 = E[X^2] = \int_{-\infty}^{+\infty} x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx =$$

$$= \int_{-\infty}^{+\infty} \underbrace{x}_{f(x)} \underbrace{\left(\frac{1}{\sqrt{2\pi}} x e^{-\frac{x^2}{2}} \right)}_{g'(x)} dx =$$

Example

$X \sim N(60, 0.0004)$

$$Y = \frac{X - 60}{0{,}02}$$

$Y \sim N(0, 1)$

$P(59{,}95 < X < 60{,}05)$

$$P\left(\frac{-0{,}05}{0{,}02} \leq Z \leq \frac{0{,}05}{0{,}02}\right) = P(-2{,}5 \leq Z \leq 2{,}5)$$

Example

if $Z \sim N(0,1)$ what is $P(Z > 0{,}92)$

$P(Z > 0{,}92) = 1 - P(Z \leq 0{,}92) =$

$= 1 - 0{,}8212$

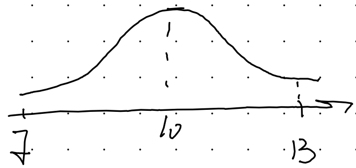
$P(-2{,}5 < Z < 2{,}5) = \Phi(2{,}5) - \Phi(-2{,}5) =$

$= \Phi(2{,}5) - (1 - \Phi(2{,}5)) = 2\Phi(2{,}5) - 1 =$

$= 2 \cdot 0{,}9938 - 1 \approx 0{,}9876$

$X \sim N(10, 1)$

$P(X < 10) = 0{,}5$



$P(X < 11{,}3) = P(Z < 1{,}3) = 0{,}9032$

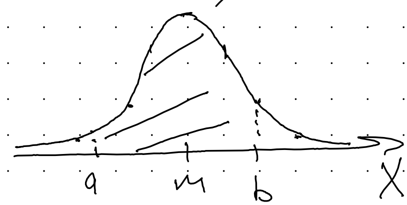
Linear combination of Normal RVs

Theorem 3: if $X \sim N(m_x, \sigma_x^2)$ and $Y \sim N(m_y, \sigma_y^2)$ and X and Y are independent then

$X + Y \sim N(m_x + m_y, \sigma_x^2 + \sigma_y^2)$

$X - Y \sim N(m_x - m_y, \sigma_x^2 + \sigma_y^2)$

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



$P(a < X < b) =$

$= P\left(\frac{a-m}{\sigma} < \frac{X-m}{\sigma} < \frac{b-m}{\sigma}\right) =$

$= P\left(\frac{a-m}{\sigma} < Y < \frac{b-m}{\sigma}\right) = \Phi\left(\frac{b-m}{\sigma}\right) - \Phi\left(\frac{a-m}{\sigma}\right)$

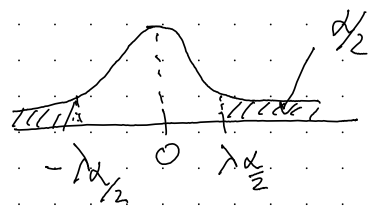
page 137

if we take

$a = m - \lambda_\alpha \cdot \sigma$

$b = m + \lambda_\alpha \sigma$

$P(m - \lambda_{\alpha/2} \cdot \sigma < X < m + \lambda_{\alpha/2} \sigma) =$



$= P(-\lambda_{\alpha/2} < Y < \lambda_{\alpha/2}) =$

Example

$$X \sim N(25, 0{,}4)$$

$$\sum_{i=1}^{20} X_i \sim N\left(\sum_{i=1}^{20} 25, \sum_{i=1}^{20} 0{,}4\right) \sim N(20\cdot 25, 20\cdot 0{,}4)$$

$$\sim N(500, 8)$$

$$P(X > 505) = 1 - P(X < 505) =$$

$$= 1 - P\left(X < \frac{5}{2\sqrt{2}}\right) \approx 1 - P(X < 1{,}77) \approx$$

$$\approx 1 - 0{,}9616 = 0{,}0384$$

Why normal distribution is useful?

- Real world experiment have normal distribution
- easy to work with (?) is it well researched?
- we can approximate to ND

The Normal distribution approx. to the binomial

The Poisson approximation to the Binomial

Application of the normal distr.

Theorem (Central Limit Theorem)
if $X_1, X_2, \dots$ is an inf sequence of independent identically distrib. rv's, each with expectation m and stand deviation $\sigma > 0$, and if we set $Y_n = X_1 + X_2 + \dots + X_n$
then we have

$$P\left(a < \frac{Y_n - nm}{\sigma\sqrt{n}} < b\right) \to \Phi(b) - \Phi(a)$$

as $n \to \infty$

$E[Y_n] = nm$ by properties

$DY_n = \sqrt{n}\,\sigma$

Def: if Z_n, $n = 1, 2, \dots$ is inf sequence of RVs and it is possible to to find number A_n and B_n $n = 1, 2, \dots$ such that

$$P\left(a < \frac{Z_n - A_n}{\sqrt{B_n}} < b\right) \to \Phi(b) - \Phi(a)$$ when $n \to \infty$

then Z_n is said to be asymptotically normally distributed with parameters A_n and B_n

$$Z_n \sim AsN(A_n, B_n)$$

Using the def the message of Theorem 4 is that $Y_n \sim AsN(nm, n\sigma^2)$

Corollary

if $X_1, X_2, \ldots$ is inf seq of indep. identically distributed RVs with m and σ then

$$\overline{X}_n \sim AsN(m, \sigma^2/n)$$