Department of Mathematics
College of William & Mary

MATH 451/551

Chapter 5. Common Continuous Distribution

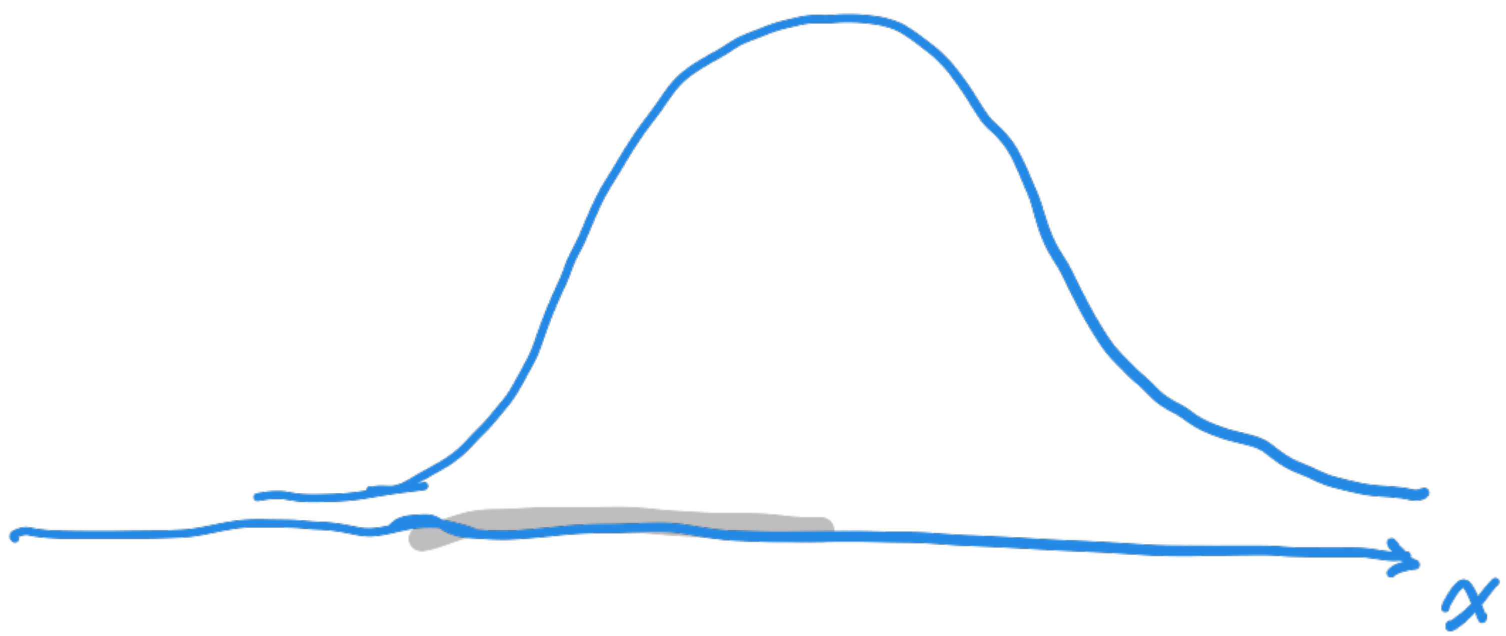
5.5 Normal Distribution

GuanNan Wang
gwang01@wm.edu

Motivating Examples

Motivating Examples

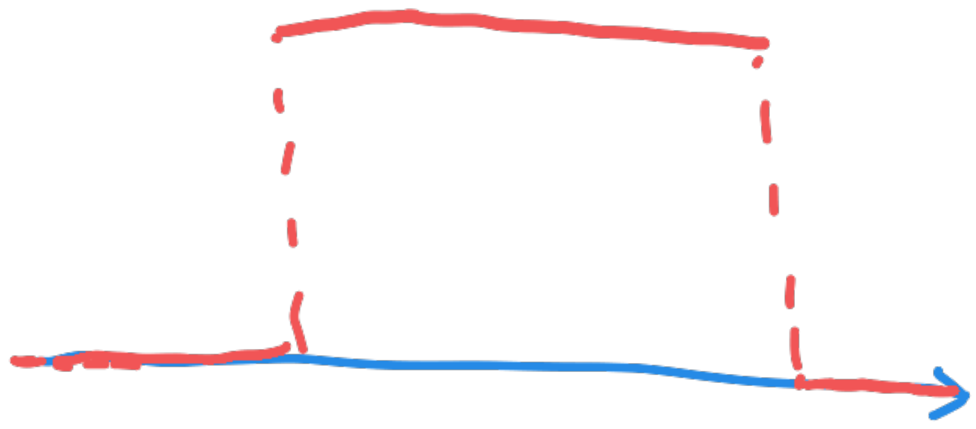
- Many random variables that arise in practice have bell-shaped distributions. Examples include
  - cholesterol levels of 50-year-old men
  - heights of adult women
  - weights of newborn babies
  - crop yields
  - ball bearing diameters



potential Bell-shaped pdf

$$f^*(x) = e^{-\frac{x^2}{2}} \qquad x \in R$$

$$\Rightarrow a = \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}}\, dx = 2\int_0^{\infty} e^{-\boxed{\frac{x^2}{2}}}\, dx$$



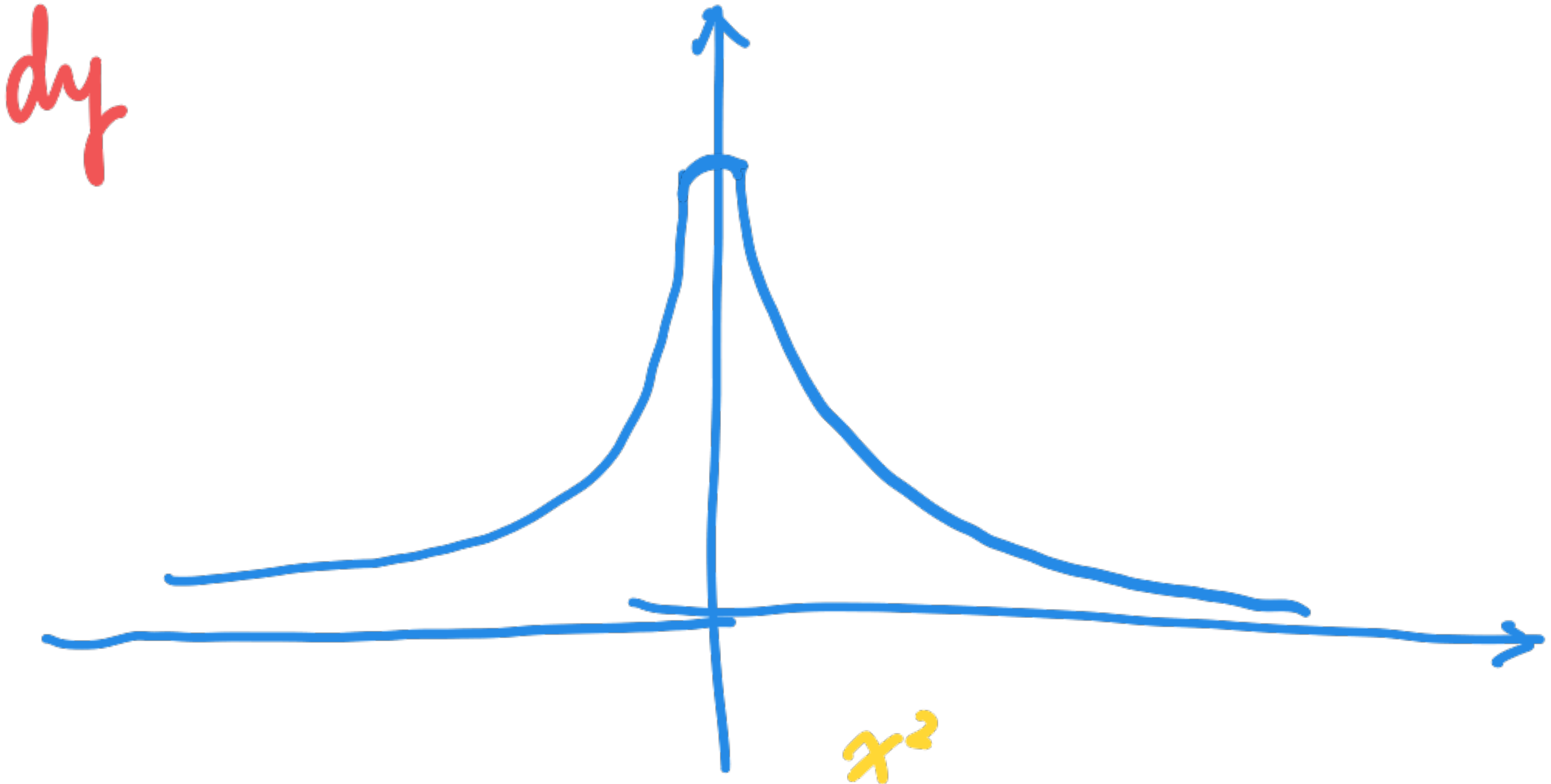
Let $y = \frac{x^2}{2}$ then $x = \sqrt{2y}$

$y \geq 0$ $\qquad dx = \frac{1}{\sqrt{2y}}\, dy$

$$\overset{y=\frac{x^2}{2}}{=} 2\int_0^{\infty} e^{-y} \frac{1}{\sqrt{2y}}\, dy$$

$$= \sqrt{2} \int_0^{\infty} y^{-\frac{1}{2}} e^{-y}\, dy$$

$$= \sqrt{2}\, \Gamma\left(\tfrac{1}{2}\right) = \sqrt{2\pi}$$



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

Standard Normal Distribution

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

normal

$$X \sim N(\mu = 0, \sigma^2 = 1)$$

$$E(X) = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 0$$

$$V(X) = E(X^2) - \{E(X)\}^2 = E(X^2)$$

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

Let $y = \frac{x^2}{2}$ then $x = \sqrt{2y}$

$$dx = \frac{1}{\sqrt{2y}} dy$$

$$\overset{y=\frac{x^2}{2}}{=} 2\int_0^{\infty} x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 2\int_0^{\infty} 2y \frac{1}{\sqrt{2\pi}} e^{-y} \frac{1}{\sqrt{2y}} dy$$

$$= \frac{2}{\sqrt{\pi}} \int_0^{\infty} y^{\frac{1}{2}} e^{-y} dy = \frac{2 \, \Gamma(1+\frac{1}{2})}{\sqrt{\pi}} = \frac{2 \cdot \frac{1}{2} \cancel{\Gamma(\frac{1}{2})}}{\cancel{\sqrt{\pi}}}$$

$$= 1$$

$$\underbrace{z = \frac{x-\mu}{\delta}} \Rightarrow \underbrace{x = \mu + \delta z} \qquad dz = \frac{1}{\delta}dx$$

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

$$\int_{-\infty}^{\infty} \underbrace{\frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\delta^2}} \frac{1}{\delta}} \, dx = 1$$

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

$$\underbrace{X \sim N(\mu, \delta^2)}$$

$$E(X) = E(\mu + \delta z) = \underbrace{E(\mu)} + E(\delta z)$$

$$= \mu + \delta \underbrace{E(z)}_{0} = \mu$$

$$V(X) = V(\mu + \delta z) = V(\delta z) = \delta^2 \underbrace{V(z)}_{1}$$

$$= \delta^2$$

Normal Distribution

Normal Distribution

A continuous random variable X with probability density function

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

for some real constants μ and $\sigma > 0$ is a $N(\mu, \sigma^2)$ random variable.

MGF

$$M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi}\,\delta} e^{-\frac{(x-\mu)^2}{2\delta^2}} \, dx$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\,\delta} \exp\left\{ \frac{-1}{2\delta^2} \left(-2\delta^2 x t + (x-\mu)^2 \right) \right\}$$

$$= \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\,\delta} \exp\left\{ -\frac{(x-(\mu+t\delta^2))^2}{2\delta^2} \right\} dx}_{1} \cdot \exp\left\{ \mu t + \frac{1}{2} t^2 \delta^2 \right\}$$

$$= \exp\left\{ \mu t + \frac{1}{2} t^2 \delta^2 \right\}, \quad t \in R.$$

Mean

$$E(X) = \frac{dM_X(t)}{dt}\Big|_{t=0} = \frac{d\, e^{\mu t + \frac{1}{2}\delta^2 t^2}}{dt}\Big|_{t=0}$$

$$= \underbrace{\Big(\mu + \underbrace{\frac{2\delta^2 t}{2}}_{0}\Big)\underbrace{e^{\mu t + \frac{1}{2}\delta^2 t^2}}_{1}}\Big|_{t=0} = \mu$$

$$E(X^2) = \frac{d^2 M_X(t)}{dt^2}\Big|_{t=0} = \delta^2 \underbrace{e^{\mu t + \frac{1}{2}\delta^2 t^2}}_{1} + (\mu + \underbrace{\delta^2 t}_{0})^2 \underbrace{e^{\mu t + \frac{1}{2}\delta^2 t^2}}_{1}\Big|_{t=0}$$

$$= \delta^2 + \mu^2$$

$$V(X) = E(X^2) - \{E(X)\}^2 = \delta^2 + \mu^2 - \mu^2 = \delta^2$$

Variance

R Functions

Function	Returned Value
dnorm(x, μ, σ)	calculates the probability density function $f(x)$
pnorm(x, μ, σ)	calculates the cumulative distribution function $F(x)$
qnorm(u, μ, σ)	calculates the percentile (quantile) $F^{-1}(u)$
rnorm(m, μ, σ)	generates m random variates

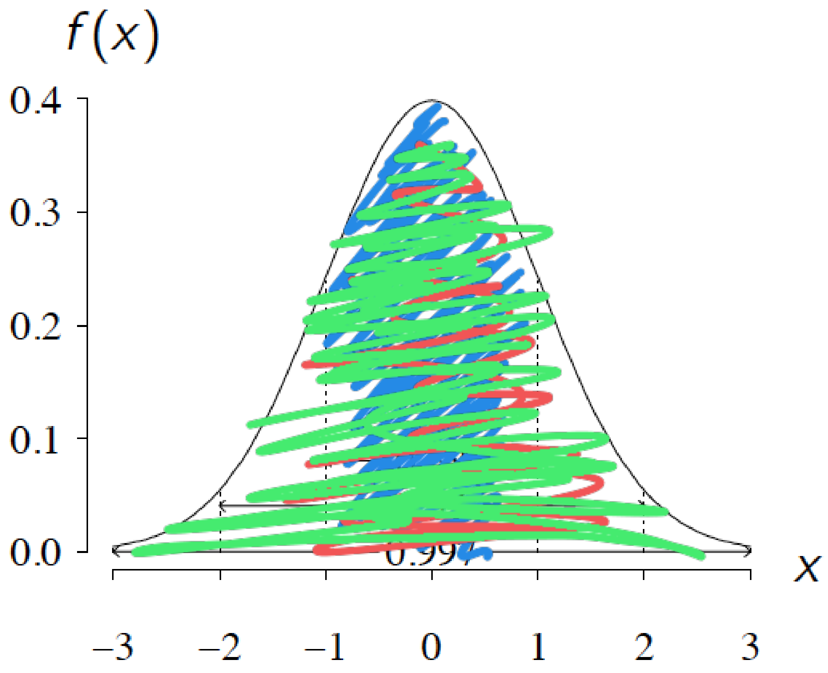
Properties

normal

- Shorthand: $X \sim N(\mu, \sigma^2)$
- Since $f(x)$ symmetric, mean, median, and mode all equal μ
- The pdf has inflection points at $\mu \pm \sigma$
- Special case: standard normal distribution $Z \sim N(0, 1)$
- The cumulative distribution function of $X \sim N(\mu, \sigma^2)$ is

$$\int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(w-\mu)^2}{2\sigma^2}}\, dw, \quad -\infty < x < \infty.$$

- Some authors use $\Phi(x)$, rather than $F(x)$, for the standard normal cdf $\quad \mu\pm\sigma \quad \mu\pm2\sigma \quad \mu\pm3\sigma$
- Empirical Rule (68–95–99.7 Rule)



Theorem 5.3

If $X \sim N(\mu, \sigma^2)$ then $Y = aX + b \sim N(a\mu + b,\ a^2\sigma^2)$ for real-valued constants $a \neq 0$ and b.

Proof: $M_Y(t) = E(e^{tY}) = E\left(e^{t(aX+b)}\right)$

$= E\left(e^{atX} \cdot e^{bt}\right) = e^{bt} E\left(e^{atX}\right)$

$= e^{bt} M_X(at) = e^{bt} e^{\mu(at) + \frac{1}{2}\sigma^2(at)^2}$

$= \exp\left\{(a\mu + b)t + \frac{1}{2}(a\sigma)^2 t^2\right\}$

$Y \sim N\left(a\mu + b, (a\sigma)^2\right)$

Theorem 5.4

If $X \sim N(\mu, \sigma^2)$ then $Z = \frac{X-\mu}{\sigma} \sim N(0,\ 1)$.

$\frac{1}{\sigma}X - \frac{\mu}{\sigma}$

Proof: $M_Z(t) = E(e^{tZ}) = E\left(e^{t\frac{X-\mu}{\sigma}}\right)$

$= e^{-\frac{\mu}{\sigma}t} E\left(e^{\frac{t}{\sigma}X}\right) = e^{-\frac{\mu}{\sigma}t} M_X(t/\sigma)$

$= e^{\cancel{-\frac{\mu}{\sigma}t}} e^{\cancel{\frac{\mu}{\sigma}t} + \frac{1}{2}\sigma^2\frac{t^2}{\sigma^2}}$

$= e^{0\cdot t + \frac{1}{2}\cdot 1 \cdot t^2} \Rightarrow Z \sim N(0, 1^2)$

Theorem 5.5

If $X \sim N(\mu, \sigma^2)$ then $Y = \left(\underbrace{\frac{X-\mu}{\sigma}}_{}\right)^2 \sim \chi^2(1)$.

Z

$$F_Y(y) = P(Y \le y) = P(Z^2 \le y) = P(-\sqrt{y} \le Z \le \sqrt{y})$$

$$= \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = 2\int_0^{\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$$

$$\underset{w=z^2}{=} \cancel{2}\int_0^y \frac{1}{\sqrt{2\pi}} e^{-\frac{w}{2}} \frac{1}{\cancel{2}\sqrt{w}} dw$$

$w = z^2$

$\Rightarrow z = \sqrt{w}$

$dz = \frac{1}{2\sqrt{w}} dw$

$$= \int_0^y \frac{1}{\sqrt{2}\sqrt{\pi}\sqrt{w}} e^{-\frac{w}{2}} dw$$

$$= \int_0^y \frac{(\frac{1}{2})^{\frac{1}{2}}}{\Gamma(\frac{1}{2})} w^{\frac{1}{2}-1} e^{-\frac{w}{2}} dw$$

$$f_Y(y) = \frac{(\frac{1}{2})^{\frac{1}{2}}}{\Gamma(\frac{1}{2})} y^{\frac{1}{2}-1} e^{-\frac{y}{2}}$$

$Y \sim \text{Gamma}(\frac{1}{2}, \frac{1}{2})$

$\sim \chi^2_{(1)}$

Example 1

IQ scores are normally distributed with population mean 100 and population standard deviation 15.

1. Find the probability that an IQ score is less than 130.
2. Find the probability that an IQ score falls between 80 and 110.
3. Find the 99th percentile of the distribution of IQ scores.

Let X = IQ scores. $X \sim N(100, 15^2)$

$P(X < 130) = P(X \leq 130) = F_X(130) = \text{pnorm}(130, 100, 15) = 0.9772$

$P(80 < X < 110) = P(80 < X \leq 110) = F_X(110) - F_X(80) = 0.7475 - 0.0912$
$= 0.6563$

$P(X \leq x) = 99\% \Rightarrow \text{qnorm}(0.99, 100, 15) = 134.8952$

Example 2

If $X \sim N(\mu, \sigma^2)$, find $E\left(|X-\mu|\right)$.

$$E(|X-\mu|) = \int_{-\infty}^{\infty} |x-\mu| \frac{1}{\sqrt{2\pi}\delta} \exp\left\{-\frac{(x-\mu)^2}{2\delta^2}\right\} dx$$

$$= 2\int_{\mu}^{\infty} (x-\mu) \frac{1}{\sqrt{2\pi}\delta} \exp\left\{-\frac{(x-\mu)^2}{2\delta^2}\right\} dx$$

$$= \frac{2\delta}{\sqrt{2\pi}} \left[-\exp\left\{-\frac{(x-\mu)^2}{2\delta^2}\right\}\right]\Bigg|_{\mu}^{\infty}$$

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

THANK YOU!