# Continuous distributions

Five important continuous distributions:

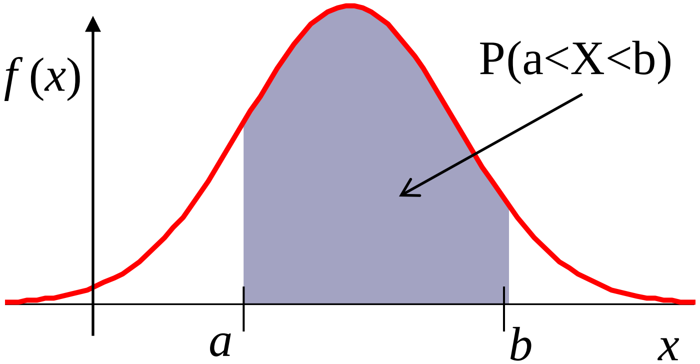
1. Uniform distribution (continuous)
2. Normal distribution
3. $\chi^2$ –distribution ["chi-square"]
4. t-distribution
5. F-distribution

# A reminder

**Definition:**
Let $X: S \to R$ be a continuous random variable.

A density function for $X$, $f(x)$, is defined by:

1. $f(x) \geq 0$ for all $x$
2. $\int_{-\infty}^{\infty} f(x)\, dx = 1$
3. $P(a < X < b) = \int_a^b f(x)\, dx$

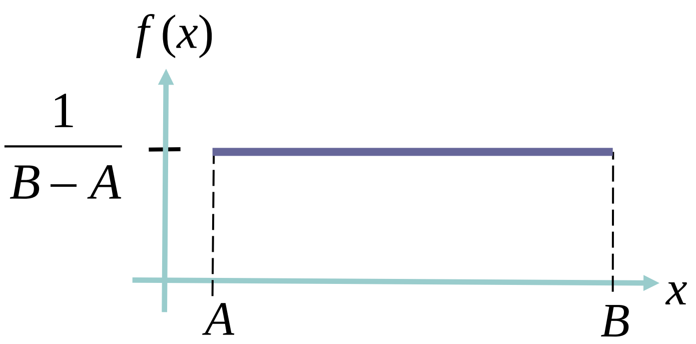


# Uniform distribution
## Definition

**Definition:**
Let X be a random variable. If the density function is given by

$$f(x) = \frac{1}{B-A} \qquad A \le x \le B$$

then the distribution of $X$ is the (continuous) uniform distribution on the interval $[A,B]$.



# Uniform distribution
## Mean & variance

**Theorem:**
Let $X$ be uniformly distributed on the interval $[A,B]$.
Then we have:

- mean of $X$: $$E(X) = \frac{A+B}{2}$$

- variance of $X$: $$Var(X) = \frac{(B-A)^2}{12}$$

# Normal distribution
## Definition

**Definition:**
Let $X$ be a continuous random variable.
If the density function is given by

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

then the distribution of $X$ is called the normal distribution with parameters $\mu$ and $\sigma^2$ (known).

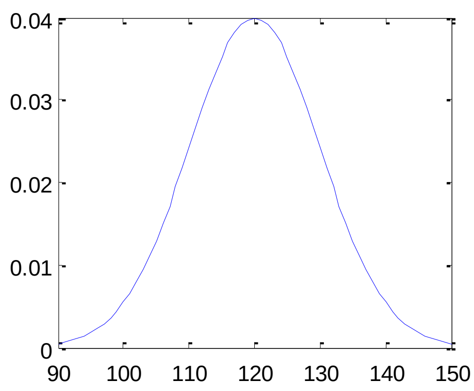
**My notation:** $X \sim N(\mu, \sigma^2)$

**The book's:** $n(x; \mu, \sigma)$ for density function

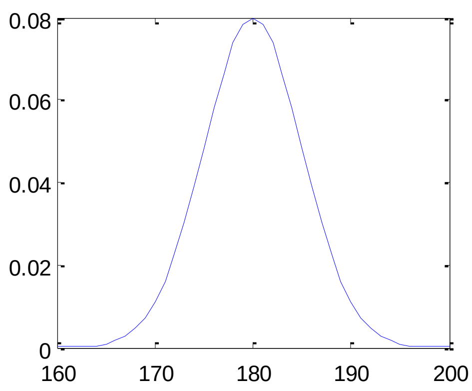
# Normal distribution
## Examples

The normal distribution is without doubt the most important continuous distribution, since many phenomena are well described by it.

IQ among AAU students




Height among AAU students




Plotting in Matlab:

```
>> x=90:1:150; y=normpdf(x,120,10); plot(x,y)
```

$\mu$ $\sigma$

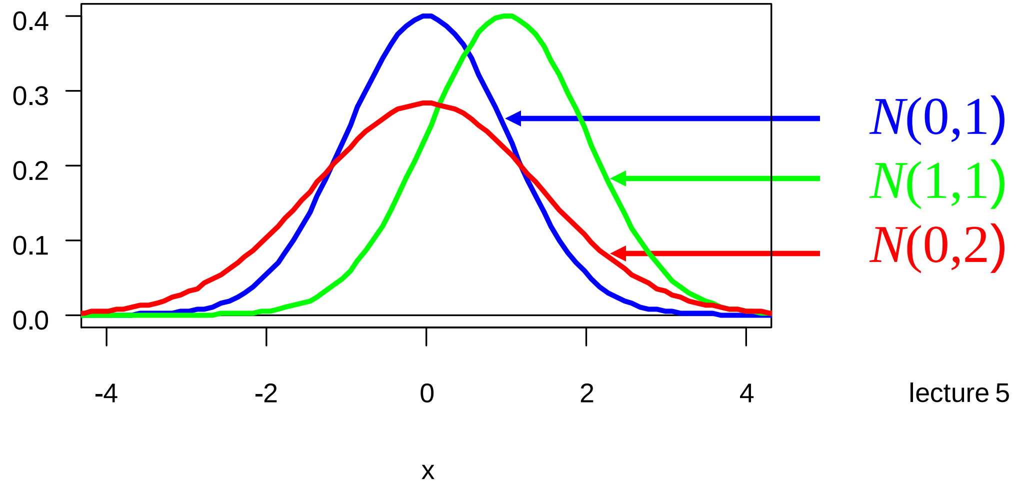
# Normal distribution
## Mean & variance

**Theorem:**

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

- mean of $X$: $E(X) = \mu$
- variance of $X$: $Var(X) = \sigma^2$

Density function:



# Normal distribution
## Standard normal distribution

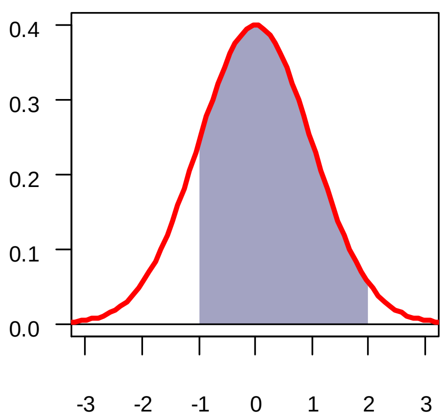
**Standard normal distribution:** $Z \sim N(0,1)$

Density function:

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

Distribution function:

$$F(z) = P(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx$$




$$P(-1 \le Z \le 2) = \int_{-1}^{2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz$$

(see Table A.3)

Notice!! Due to symmetry

$P(Z \le -z) = 1 - P(Z \le z)$

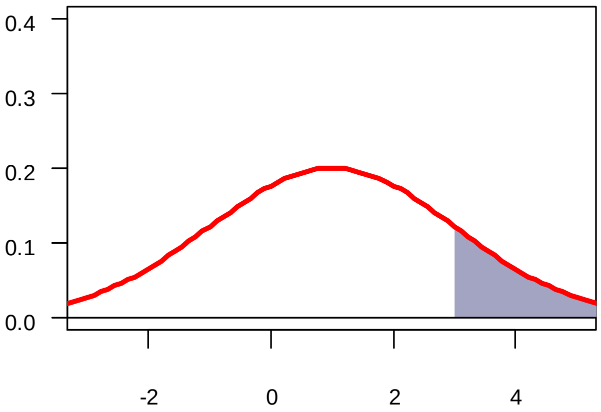
# Normal distribution
## Standard normal distribution

Standard normal distribution, N(0,1), is the only normal distribution for which the distribution function is tabulated.

We typically have $X \sim N(\mu,\sigma^2)$ where $\mu \neq 0$ and $\sigma^2 \neq 1$.

**Example:** $X \sim N(1,4)$

What is $P( X > 3 )$ ?



# Normal distribution
## Standard normal distribution

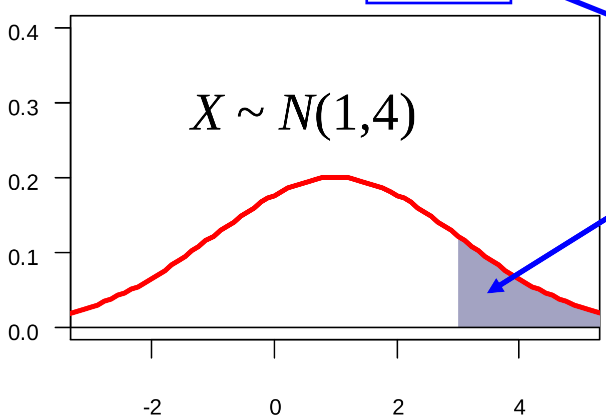
**Theorem:** Standardise

If $X \sim N(\mu,\sigma^2)$ then $\dfrac{X-\mu}{\sigma} \sim N(0,1)$

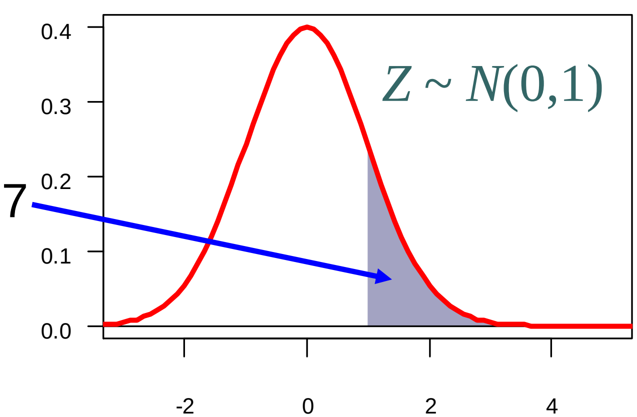
**Example cont.:** $X \sim N(1,4)$. What is $P(X>3)$ ?

$$P(X>3) = P\left(\boxed{\frac{X-1}{2}} > \frac{3-1}{2}\right) = P(Z>1) = 1 - P(Z \le 1) = 0.1587$$

$Z \sim N(0,1)$




Equal areas = 0.1587



# Normal distribution

## Standard normal distribution

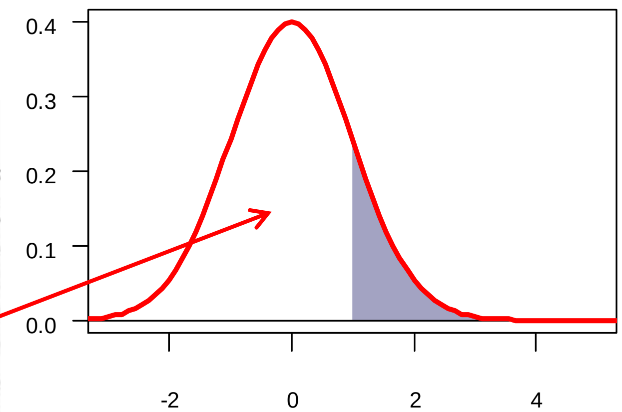



Table A.3 (continued) Areas under the Normal Curve

| z | .00 | .01 | .02 | .03 | .04 | .05 | .06 | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.0 | 0.5000 | 0.5040 | 0.5080 | 0.5120 | 0.5160 | 0.5199 | 0.5239 | 0 | | |
| 0.1 | 0.5398 | 0.5438 | 0.5478 | 0.5517 | 0.5557 | 0.5596 | 0.5636 | 0 | | |
| 0.2 | 0.5793 | 0.5832 | 0.5871 | 0.5910 | 0.5948 | 0.5987 | 0.6026 | 0 | | |
| 0.3 | 0.6179 | 0.6217 | 0.6255 | 0.6293 | 0.6331 | 0.6368 | 0.6406 | 0 | | |
| 0.4 | 0.6554 | 0.6591 | 0.6628 | 0.6664 | 0.6700 | 0.6736 | 0.6772 | 0 | | |
| 0.5 | 0.6915 | 0.6950 | 0.6985 | 0.7019 | 0.7054 | 0.7088 | 0.7123 | 0.7157 | 0.7190 | 0.7224 |
| 0.6 | 0.7257 | 0.7291 | 0.7324 | 0.7357 | 0.7389 | 0.7422 | 0.7454 | 0.7486 | 0.7517 | 0.7549 |
| 0.7 | 0.7580 | 0.7611 | 0.7642 | 0.7673 | 0.7704 | 0.7734 | 0.7764 | 0.7794 | 0.7823 | 0.7852 |
| 0.8 | 0.7881 | 0.7910 | 0.7939 | 0.7967 | 0.7995 | 0.8023 | 0.8051 | 0.8078 | 0.8106 | 0.8133 |
| 0.9 | 0.8159 | 0.8186 | 0.8212 | 0.8238 | 0.8264 | 0.8289 | 0.8315 | 0.8340 | 0.8365 | 0.8389 |
| 1.0 | 0.8413 | 0.8438 | 0.8461 | 0.8485 | 0.8508 | 0.8531 | 0.8554 | 0.8577 | 0.8599 | 0.8621 |
| 1.1 | 0.8643 | 0.8665 | 0.8686 | 0.8708 | 0.8729 | 0.8749 | 0.8770 | 0.8790 | 0.8810 | 0.8830 |
| 1.2 | 0.8849 | 0.8869 | 0.8888 | 0.8907 | 0.8925 | 0.8944 | 0.8962 | 0.8980 | 0.8997 | 0.9015 |
| 1.3 | 0.9032 | 0.9049 | 0.9066 | 0.9082 | 0.9099 | 0.9115 | 0.9131 | 0.9147 | 0.9162 | 0.9177 |
| 1.4 | 0.9192 | 0.9207 | 0.9222 | 0.9236 | 0.9251 | 0.9265 | 0.9279 | 0.9292 | 0.9306 | 0.9319 |

$$P(Z \le 1) = 0.8413 \quad \Rightarrow \quad P(X > 3) = 1 - P(Z \le 1) = 1 - 0.8413 = 0.1587$$

# Normal distribution
## Standard normal distribution

**Example cont.:** $X \sim N(1,4)$. What is $P(X > 3)$ ?

$$P(X > 3) = 1 - P(X \leq 3) = 0.1587$$

Solution in Matlab:

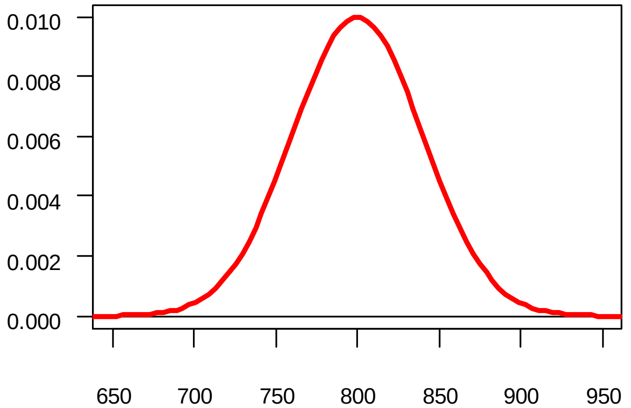
```
>> 1 - normcdf(3,1,2)
ans =
      0.1587
```

Cumulative distribution function

`normcdf(x,μ,σ)`

# Normal distribution
## Example

**Problem:**
The lifetime of a light bulb is normal distributed with mean 800 hours and standard deviation 40 hours:




1. Find the probability that the lifetime of a bulb is between 750-850 hours.
2. Find the number of hours b, such that the probability of a bulb having a lifetime longer than b is 90%.
3. Find a time period *symmetric* around the mean so that the probability of a lifetime in this interval has probability 95%

# Normal distribution
## Example

Solution in Matlab:

1. P(750 < X < 850) = ?

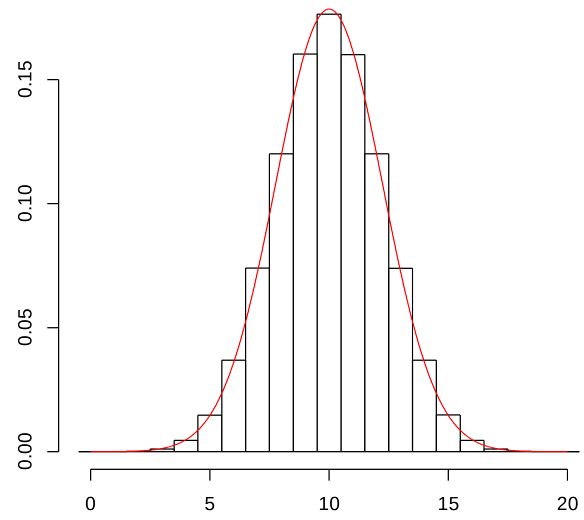
   ```
   >> normcdf(850,800,40)-normcdf(750,800,40)
   ```
   or
   ```
   >> diff(normcdf([750 850],800,40))
   ```
2. ?
3. ...

# Normal distribution
## Relation to the binomial distribution

If $X$ is binomial distributed with parameters $n$ and $p$, then

$$Z = \frac{X - np}{\sqrt{np(1-p)}}$$

is approximately normal distributed.




**Rule of thumb:**
If $np>5$ and $n(1-p)>5$, then the approximation is good

# Normal distribution

## Linear combinations

**Theorem:** linear combinations

If $X_1$, $X_2$,..., $X_n$ are independent random variables, where

$$X_i \sim N(\mu_i, \sigma_i^2), \text{ for } i = 1,2,\ldots,n,$$

and $a_1, a_2, \ldots, a_n$ are constant, then the linear combination

$$Y = a_1 X_1 + a_2 X_2 + \cdots + a_n X_n \sim N(\mu_Y, \sigma_Y^{\ 2}),$$

where

$$\mu_Y = a_1\mu_1 + a_2\mu_2 + \cdots + a_n\mu_n$$

$$\sigma_Y^2 = a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \cdots + a_n^2\sigma_n^2$$

# The $\chi^2$ distribution
## Definition

**Definition:** (alternative to Walpole, Myers, Myers & Ye)
If $Z_1, Z_2, \ldots, Z_n$ are independent random variables, where

$$Z_i \sim N(0,1), \qquad \text{for } i = 1,2,\ldots,n,$$

then the distribution of

$$Y = Z_1^{\,2} + Z_2^{\,2} + \cdots + Z_n^{\,2} = \sum_{i=1}^{n} Z_i^{\,2}$$

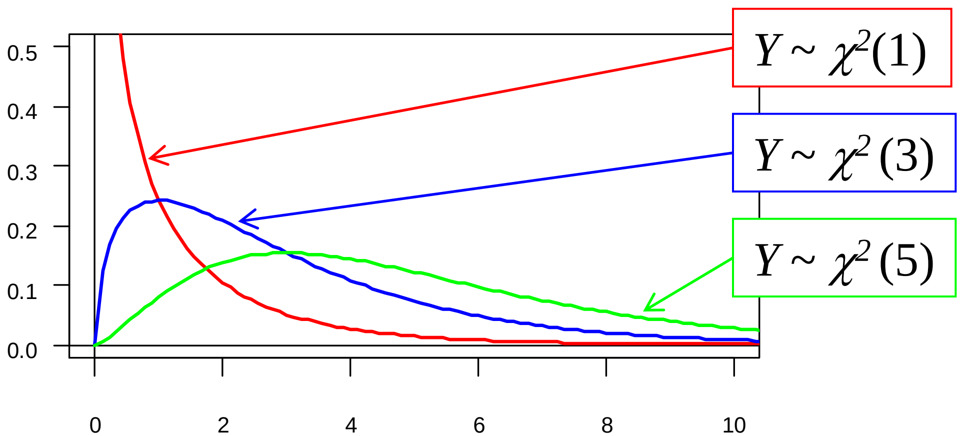
is the $\chi^2$-distribution with $n$ degrees of freedom.

**Notation:** $Y \sim \chi^2(n)$ Critical values: Table A.5

# The $\chi^2$ distribution
## Definition

**Definition:** A continuous random variable $X$ follows a $\chi^2$-distribution with $n$ degrees of freedom if it has density function

$$f(x) = \frac{1}{2^{n/2}\Gamma(n/2)} x^{n/2-1} e^{-x/2} \quad \text{for } x > 0$$




Assume $Y \sim \chi^2(n)$

$E(Y) = n$
$Var(Y) = 2n$

$E(Y/n) = 1$
$Var(Y/n) = 2/n$

# *t*-distribution
## Definition

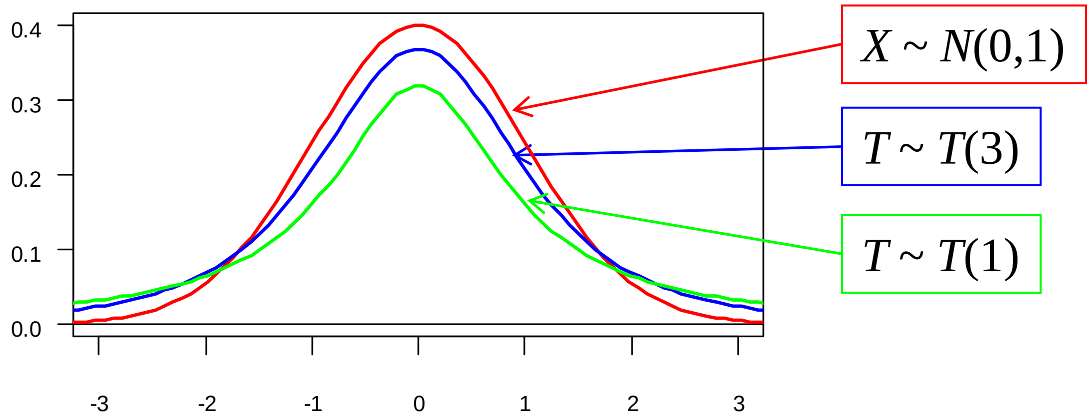
**Definition:**
Let $Z \sim N(0,1)$ and $V \sim \chi^2(n)$ be two independent random variables. Then the distribution of

$$T = \frac{Z}{\sqrt{V/n}}$$

is called the *t*-distribution with $n$ degrees of freedom.

**Notation:** $\mathrm{T} \sim t(n)$ Critical values: Table A.4

# *t*-distribution
## Compared to standard normal




- The *t*-distribution is symmetric around 0
- The *t*-distribution is more flat than the standard normal
- The more degrees of freedom the more the *t*-distribution looks like a standard normal

# *F*-distribution
## Definition

**Definition:**
Let $U \sim \chi^2(n_1)$ and $V \sim \chi^2(n_2)$ be two independent random variables. Then the distribution of

$$F = \frac{U/n_1}{V/n_2}$$

is called the *F*-distribution with $n_1$ (numerator) and $n_2$ (denominator) degrees of freedom.

**Notation:** $F \sim F(n_1\,,\, n_2)$ Critical values: Table A.6

# *F*-distribution

## Example