Probability, distribution, density

Events (samples, outcomes)

ω = an event (or sample, or outcome)

Ex: Tossing a coin, $\omega = {\rm heads}$ or $\omega = {\rm tails}$

\omega = {\rm heads} \omega = {\rm tails}

///

Ω = set of all events (samples, outcomes), called "sample space" (statistical theory) or "outcome space" (probability theory).

(Xiu 2010 p.9, Shao 2007 p.1)

Xiu 2010, Numerical methods for stochastic computations: A spectral method approach.

Shao 2007, Mathematical statistics, 2nd ed.

Ex: Tossing a coin, $\Omega = \{{\rm heads}, {\rm tails}\}$

\Omega = \{{\rm heads} , {\rm tails} \}

Thus $\omega \in \Omega$ means the event ω is either

\omega \in \Omega

"heads" (obverse) or "tails" (reverse). ///

http://en.wikipedia.org/wiki/Coin

$\mathcal{F}$ = collection of all subsets of Ω \mathcal F

Ex: $\mathcal{F} = \{\emptyset, \mathrm{heads}, \mathrm{tails}, \Omega\}$ ///

\mathcal F = \{ \emptyset , {\rm heads} , {\rm tails} , \Omega \}

Note: $\mathcal{F}$ is called a "sigma-field" or "sigma-algebra"

written as σ-field or σ-algebra

\usepackage{color}
\color{blue}{\sigma}

σ is mnemonic for "S", and "Sum", due to property

$$\bigcup_i A_i \in \mathcal{F} \, , \ \forall A_i \subset \mathcal{F} \qquad (1)$$

\displaystyle \bigcup_i A_i \in \mathcal F \ , \ \forall A_i \subset \mathcal F

(Xiu 2010 p.10, Shao 2007 p.2, for formal definition)

i.e., sum or union of any subsets of $\mathcal{F}$ is a subset of $\mathcal{F}$.

Recall the algebra on the real line $\mathbb{R}$

\usepackage{amssymb}
\mathbb R

with operations +, -, x, /, e.g., $1 + 2 \in \mathbb{R}$ ///

1 + 2 \in \mathbb R

Ex: $\emptyset \cup \Omega = \Omega \in \mathcal{F}$

\emptyset \cup \Omega = \Omega \in \mathcal F

$\emptyset \cup {\rm heads} = {\rm heads} \in \mathcal{F}$

\emptyset \cup {\rm heads} = {\rm heads} \in \mathcal F

$\Omega \cup {\rm heads} = \Omega \in \mathcal{F}$

\Omega \cup {\rm heads} = \Omega \in \mathcal F

${\rm heads} \cap {\rm tails} = \emptyset \in \mathcal{F}$

{\rm heads} \cap {\rm tails} = \emptyset \in \mathcal F

$\cap$ is also a valid operation in $\mathcal{F}$. ///

Probability of an event ω belonging to an element $A \in \mathcal F$ is a non-negative number ("measure"), A \in \mathcal F

denoted by $P(\omega \in A) = P(A)$ P (\omega \in A) = P(A)

$P : \mathcal F \rightarrow \mathbb R_0^+$ set of non-negative real numbers

P : \mathcal F \rightarrow \mathbb R_0^+

$[A \in \mathcal F] \mapsto [P(A) \in \mathbb R_0^+]$

[A \in \mathcal F] \mapsto [P(A) \in \mathbb R_0^+]

Ex: $P({\rm heads}) = P({\rm tails}) = \frac{1}{2}$

P ({\rm heads}) = P ({\rm tails}) = \frac{1}{2}

$P(\emptyset) = 0$

P (\emptyset) = 0

$\Omega = \{{\rm heads}, {\rm tails}\} = \{{\rm heads}\} \cup \{{\rm tails}\}$

\Omega = \{{\rm heads} , {\rm tails} \} = \{{\rm heads} \} \cup \{{\rm tails} \}

$P(\omega \in \Omega) = P({\rm heads}) + P({\rm tails}) = 1$

P (\omega \in \Omega) = P ({\rm heads}) + P ({\rm tails}) = 1

Probability that the event ω is either heads or tails.

///

$\cup \equiv \text{"or"}$

\cup \equiv \color{red}{\rm ``or"}

Ex: $\{\text{heads}\} \cup \{\text{tails}\} = \Omega$ heads or tails

\{{\rm heads} \} \cup \{{\rm tails} \} = \Omega

///

$\cap \equiv \text{"and"}$

\cap \equiv \color{red}{\rm ``and"}

Ex: $\{\text{heads}\} \cap \{\text{tails}\} = \emptyset$ heads and tails

\{{\rm heads} \} \cap \{{\rm tails} \} = \emptyset

///

Ex: A dice with 6 facets

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

\Omega = \{ 1,2,3,4,5,6 \}

$\{6\} \in \mathcal{F}$ $\{3, 6\} \in \mathcal{F}$

\{ 3,6 \} \in \mathcal F

$A := \{2, 3, 5\} \in \mathcal{F}$ $\{1, 3, 4, 6\} \in \mathcal{F}$

A := \{ 2,3,5 \} \in \mathcal F

\{ 1,3,4,6 \} \in \mathcal F

$$P(\omega \in A) = P(2) + P(3) + P(5)$$

$$= 3\frac{1}{6} = \frac{1}{2}$$

P (\omega \in A) = P(2) + P(3) + P(5) = 3 \frac16 = \frac12

///

Random variable   usually denoted in capital letters

X = random variable $X : \Omega \to \mathbb{R}$

X : \Omega \to \mathbb R

$$\omega \mapsto X(\omega) \qquad (1)$$

\omega \mapsto X(\omega)

$X(\omega)$ = (arbitrary) number selected to represent each event ω in Ω

Ex: Typically $X(\text{heads}) = 1$

X({\rm heads}) = 1

$$X(\text{tails}) = 0$$

X({\rm tails}) = 0

But it is also possible to select (even though not a good choice, since not as "mnemonic" as {0,1})

$$X(\text{heads}) = 5$$

X({\rm heads}) = 5

$$X(\text{tails}) = -3$$

X({\rm tails}) = -3

///

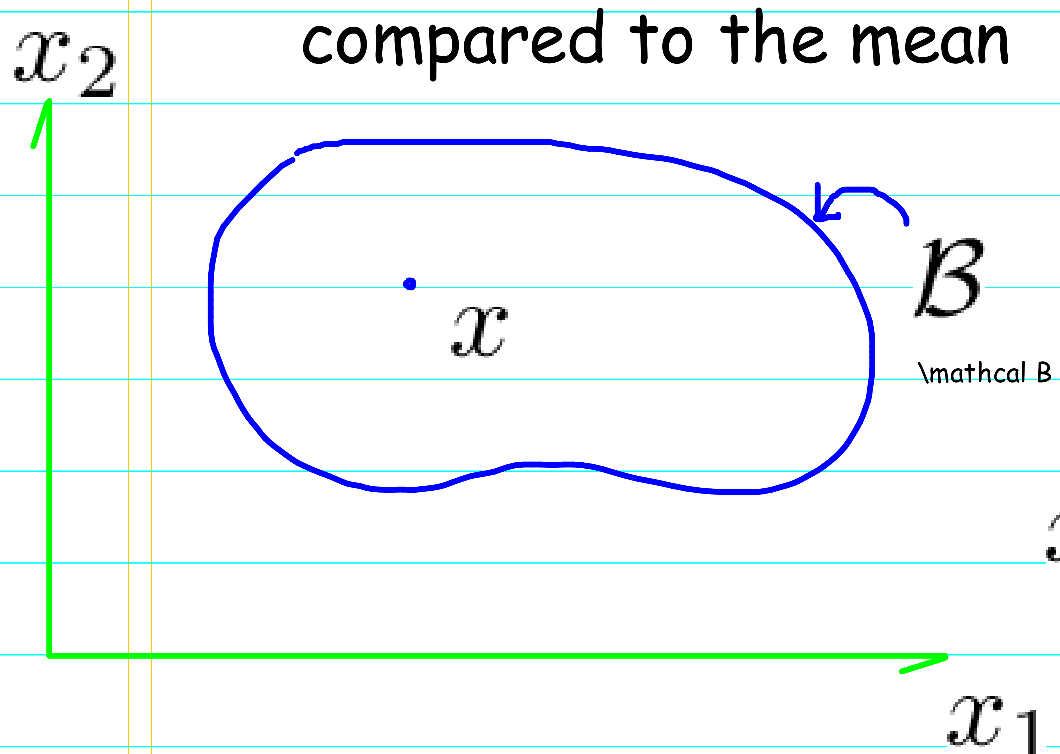
Ex: Turbulent flows

Ω can be thought of as a set of repeated experiments (samples) to verify, say, a hypothesis or observation on a given flow.

$$\Omega = \left\{ \omega_1 , \omega_2 , \cdots , \omega_{n_{exp}} \right\}$$

\Omega = \left\{ \omega_1 , \omega_2 , \cdots , \omega_{n_{exp}} \right\}

n_{exp} total number of repeated experiments, e.g., until the standard deviation is small enough compared to the mean



$$x = (x_1 , x_2 , x_3)$$

x = (x_1 , x_2 , x_3)

U_i (x,t,\omega_k)

$U_i(x,t,\omega_k)$ ith velocity component (a random variable) at (x,t) in experiment ω_k

///

Note: Technically, a random variable is the mapping ("measurable function")

$$X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}) \quad (1)$$

X : (\Omega , \mathcal F) \to (\mathbb R , \mathcal B)

$(\Omega, \mathcal{F})$ Event space Ω endowed with σ-algebra $\mathcal{F}$

$(\mathbb{R}, \mathcal{B})$ Set of real numbers $\mathbb{R}$ endowed with "Borel σ-algebra" $\mathcal{B}$ (sigma-algebra of finite open subsets of $\mathbb{R}$) (Shao 2007 p.7) ///

Ex:

$$\mathcal{B} = \sigma\Big(\{ (a,b] : a,b \in \mathbb{R} \} \Big) \quad (2)$$

\mathcal B = \sigma \Big(\{ (a,b] : a,b \in \mathbb R \} \Big)

σ ↑ σ-algebra; $(a,b]$ ↑ finite half-open interval in $\mathbb{R}$

$\{(a,b] : a,b \in \mathbb{R}\}$ Set of finite open intervals in $\mathbb{R}$

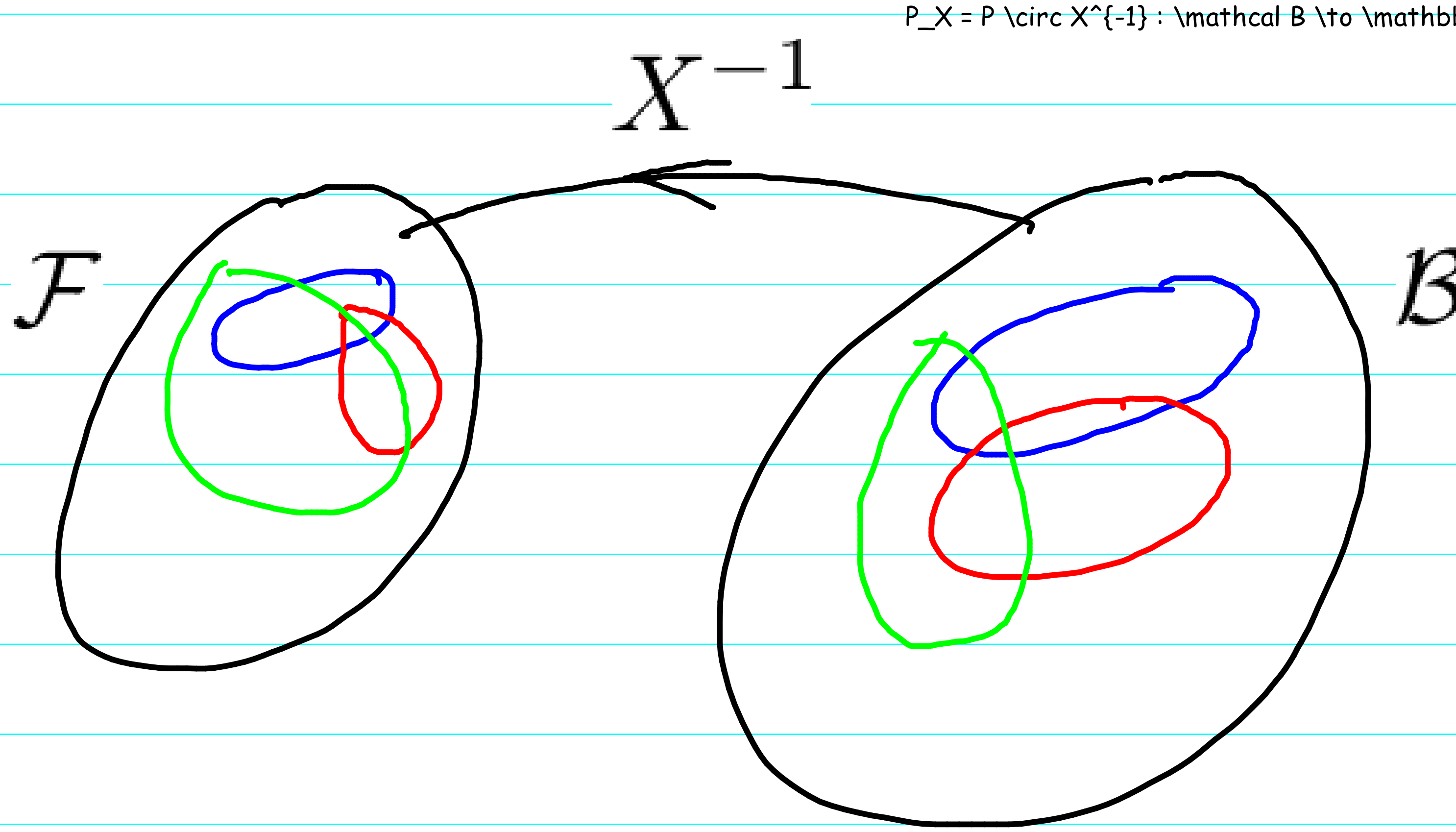
This choice of $\mathcal{B}$ allows for the probability of $X \in (a,b]$, i.e., $P(X \in (a,b])$.

(Xiu 2010 p.11) ///

P_X Law or distribution of X

$$P_X = P \circ X^{-1} : \mathcal{B} \to \mathbb{R}_0^+ \quad (1)$$

P_X = P \circ X^{-1} : \mathcal B \to \mathbb R_0^+



In practice, only to refer to an open interval in $\mathcal{B}$

$$[\underbrace{(a,b]}_{} \in \mathcal{B}] \mapsto [P_X((a,b]) \in \mathbb{R}_0^+] \quad (2)$$

half-open interval in $\mathcal{B}$

[(a,b] \in \mathcal B] \mapsto [P_X ((a,b]) \in \mathbb R_0^+]

$$P_X((a,b]) \equiv P_X(X(\omega) \in (a,b]) \quad (3)$$

P_X ((a,b]) \equiv P_X (X(\omega) \in (a,b])

i.e., the probability that $a < X(\omega) \le b$

a < X(\omega) \le b

Note:

In case events were already representable by real numbers, then the event space is already $\Omega \equiv \mathbb{R}$.

It is then not necessary to mention ω, but directly X. An example of such random variable is a velocity component in a turbulent flow.

Pope 2000 p.xx ///

Pope 2000, Turbulent flows.

$F_X(x)$ Cumulative distribution function (cdf)

$$F_X(x) := P_X((-\infty, x]) \overset{(1)}{=} P_X(X \le x)$$

F_X (x) := P_X ((-\infty , x]) = P_X (X \le x)

$f_X(x)$ Probability density function (pdf)

$$f_X(x) := \frac{d}{dx} F_X(x) \qquad (2)$$

f_X (x) := \displaystyle \frac{d}{dx} F_X (x)

$$F_X(x) = \int_{-\infty}^{x} f_X(t) dt \qquad (3)$$

F_X (x) = \displaystyle \int_{-\infty}^{x} f_X (t) dt

Note: Notation X and x

In recent literature, an uppercase letter, e.g., X , is used to designate a random variable, whereas the corresponding lowercase letter, e.g., x , is used to designate the real variable that is the upper bound of X.

Kolmogorov (Kolmo) 1933, Foundations of the theory of probability. Famous work influencing subsequent mathematical probability and statistics literature.

$$\underbrace{F^{(x)}(a)}_{\text{cdf}} = \int_{-\infty}^{a} \underbrace{f^{(x)}(a)}_{\text{pdf}} da \qquad (1)$$

Kolmo 1933 p.24

F^{(x)} (a) = \displaystyle \int_{-\infty}^{a} f^{(x)} (a) da

Modern notation, more mnemonic, is close to that chosen by Kolmo 1933.

///

Note: Terminologies

$F_X(x)$ = Cumulative distribution function (cdf)

Shao 2007 p. Pope 2000 p.

Distribution function

Kolmo 1933 p.19 Xiu 2007 p.

$f_X(x)$ = Probability density function (pdf)

Shao 2007 p. Pope 2000 p.

Probability mass function

?? 20?? p. ???? 20?? p.

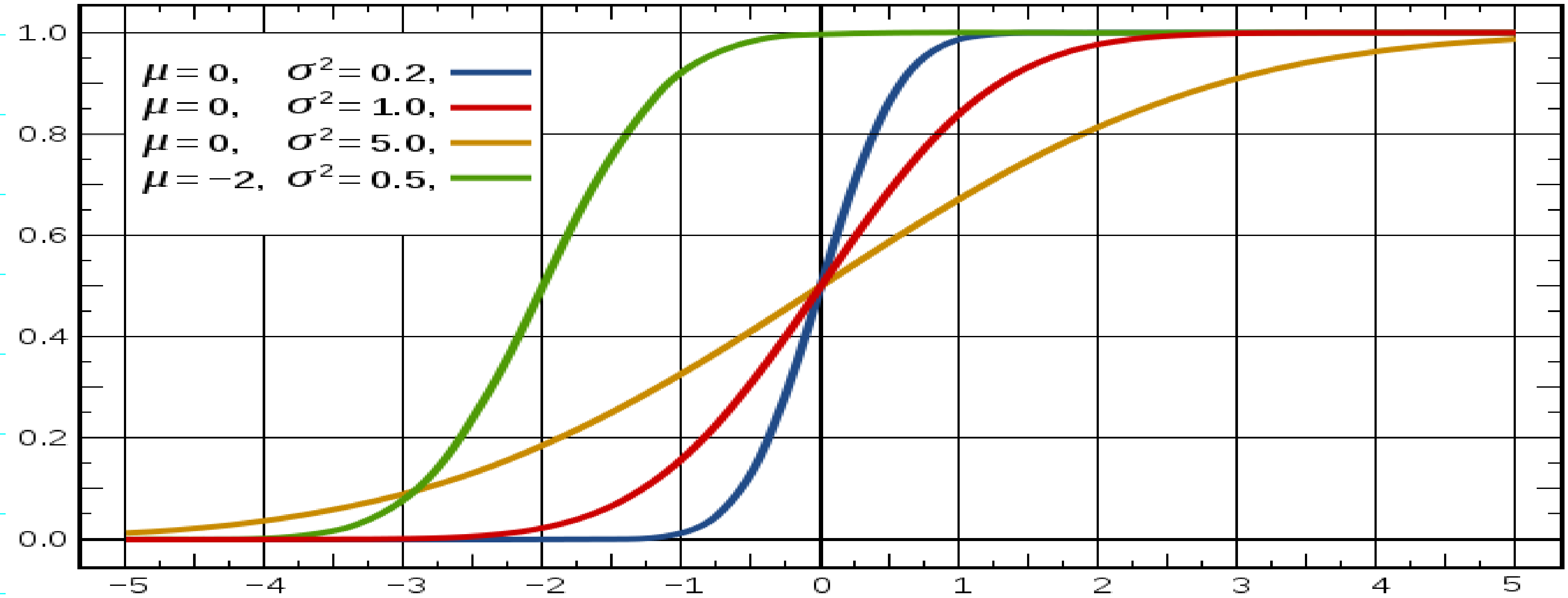
Density function

Kolmo 1933 p.19 Xiu 2007 p. ///

Normal (Gaussian) distribution

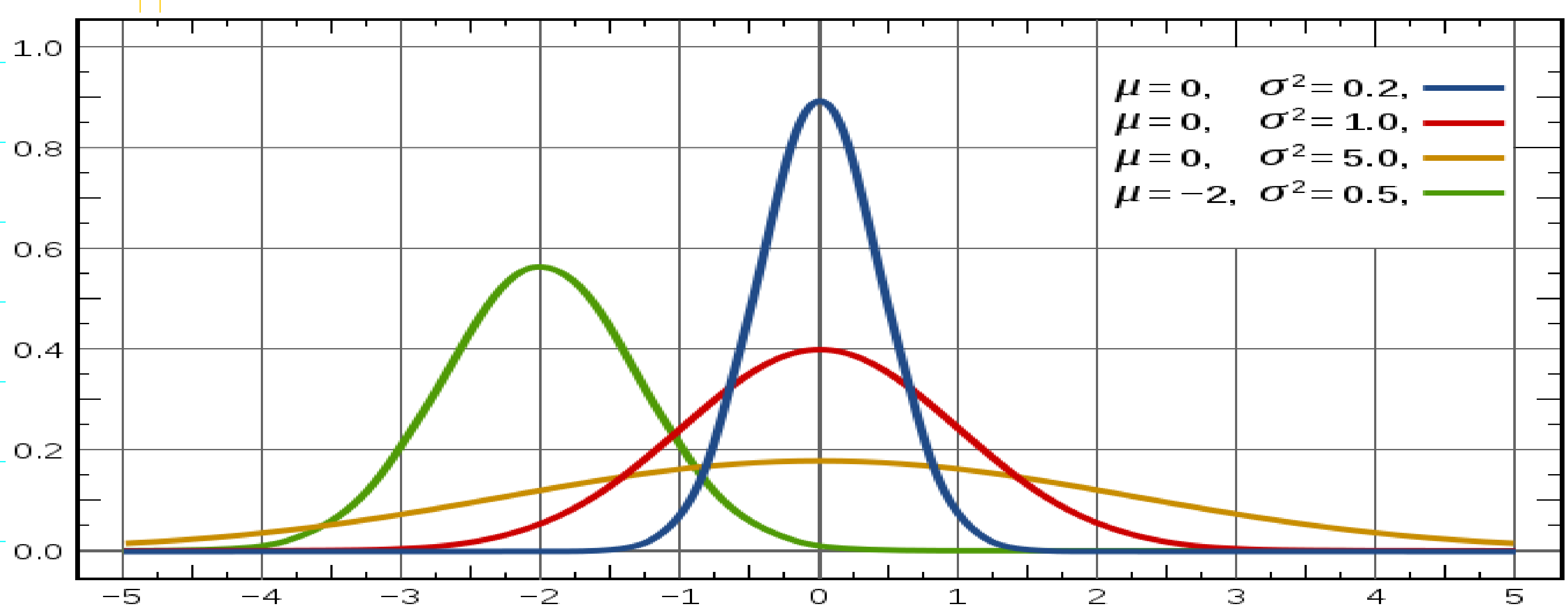
cdf $$F_X(x) = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{x}{\sqrt{2}}\right)\right] \quad (1)$$

F_X (x) = \frac12\left[\, 1 + \operatorname{erf} \left(\displaystyle \frac{x}{\sqrt{2}} \right) \right]



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

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



http://en.wikipedia.org/wiki/Normal_distribution

10 Deutsch mark bank note

Gaussian (normal) pdf

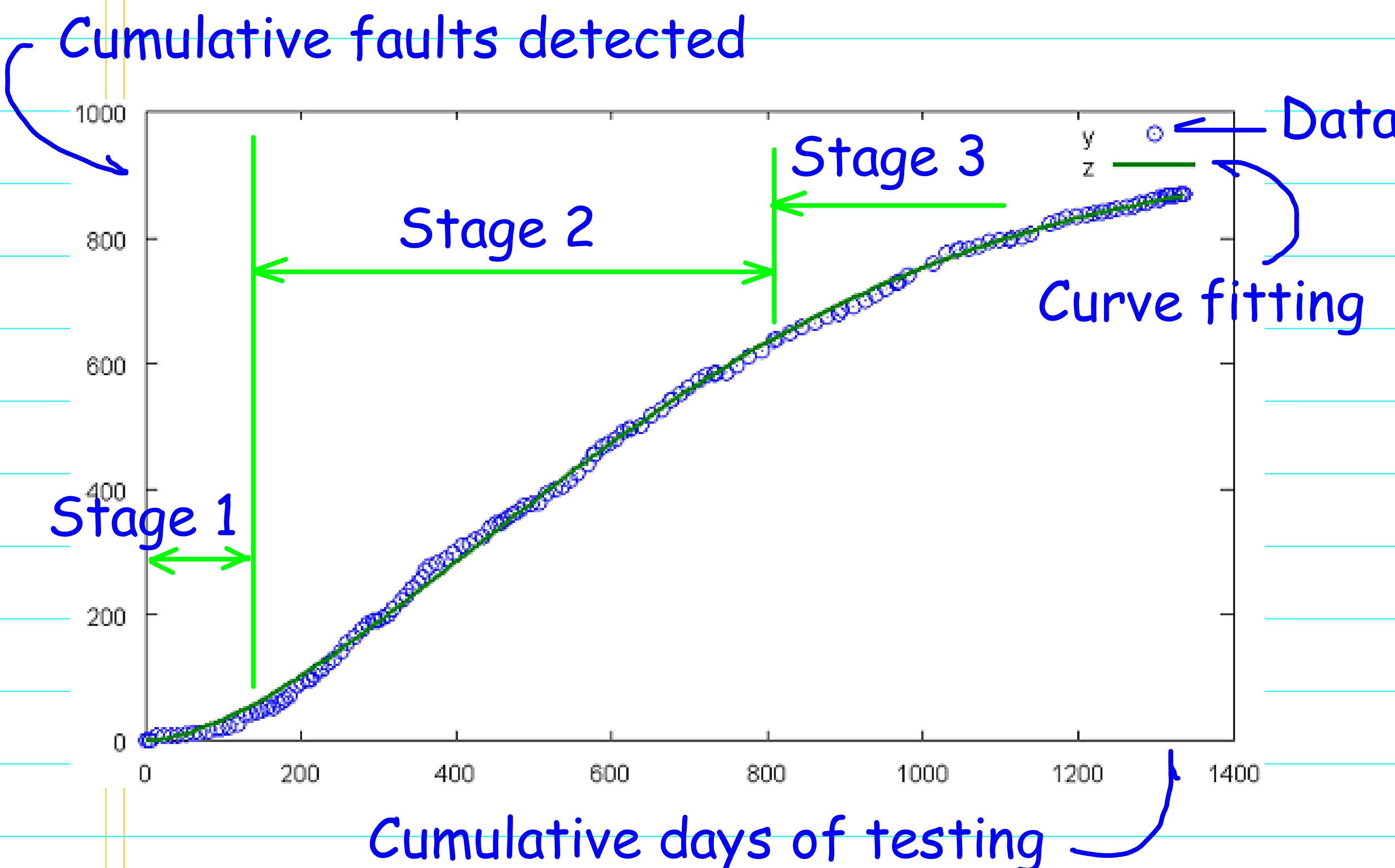
Gauss

http://en.wikipedia.org/wiki/Deutsche_Mark

http://en.wikipedia.org/wiki/Federal_Reserve_Note

Weibull distribution, exponential distribution

Application: Testing and debugging of large software (~ 7x10^6 non-commentary source lines)



Observation: Rate of fault detection (slope) slow at first, then rises up quickly, then decreases with time.

Major decision: When to stop testing ?

Important factor: Number of new faults detected if testing period were extended. Need model to predict.

Lawless & Cook 2007 p.5, p.368; Dalal & McIntosh 1994

3 stages of fault-detection rate:

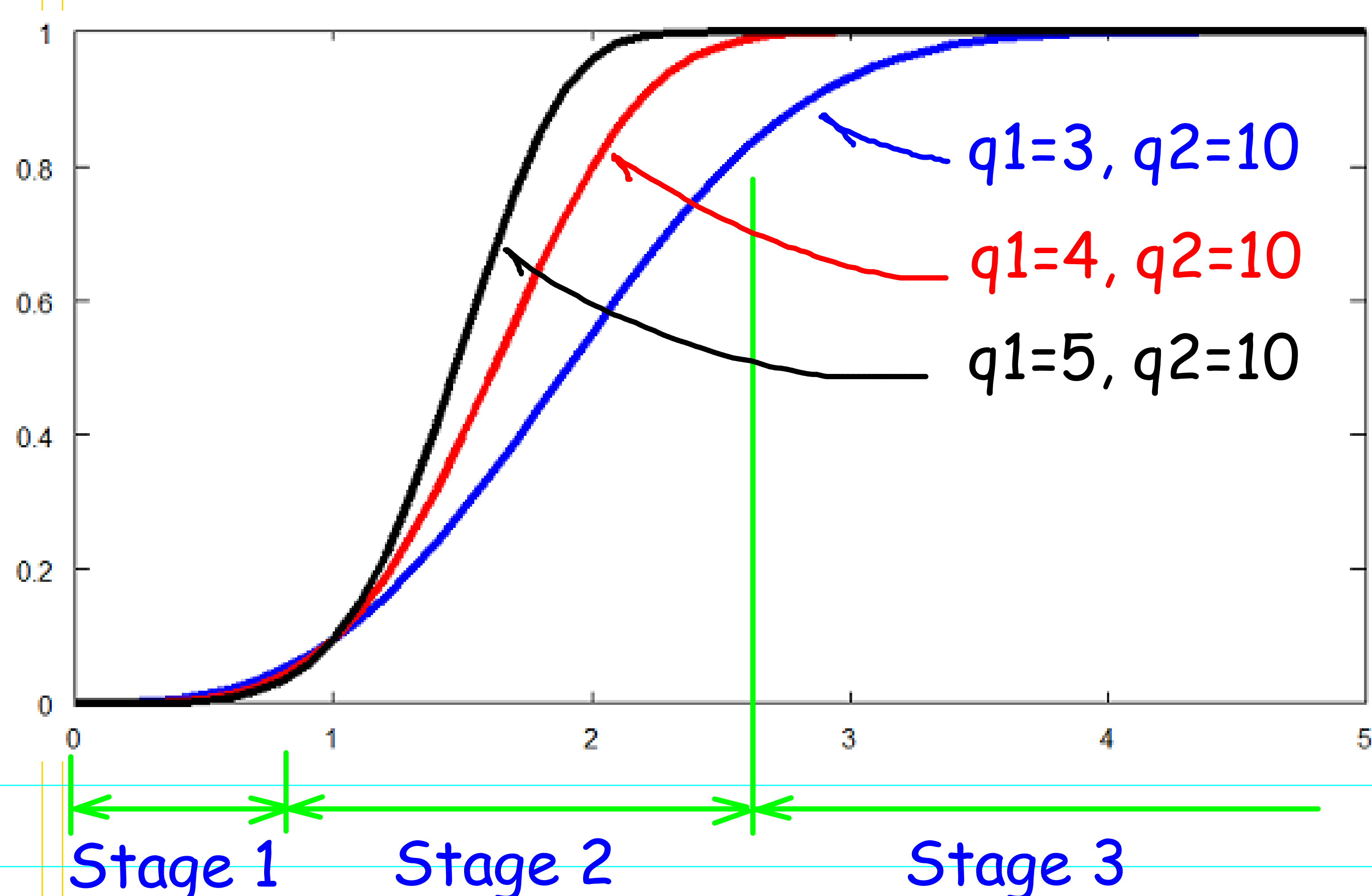
Stage 1: Low fault-detection rate at the beginning

Stage 2: Increasing and steep fault-detection rate in the middle

Stage 3: Slowing fault-detection rate at the end

$$F_X(x) = 1 - \exp\left(-\frac{x^{q_1}}{q_2}\right) \quad (1)$$

F_X (x) = 1 - {\rm exp} \left(- \displaystyle \frac{x^{q_1}}{q_2} \right)



HW x.x: Find out the meaning of the parameter q_2 in

Provide plots for visual explanation.

Take successive derivatives of $F_X(x)$ to propose logical demarcations for the different stages, i.e., Stage 1 goes from what value to what value, etc. ///

Formula to fit data points

$$z(x) = p_3 \left[1 - \exp \left(-\frac{x}{p_2} \right)^{p_1} \right] \qquad (1)$$

z(x) = p_3 \left[1 - {\rm exp} \left(- \frac{x}{p_2} \right)^{p_1} \right]

Least square curve fitting: Legendre 1805

Gauss 1795

Stiegler 1986 p.15

Stiegler 1986, The history of statistics.

Nonlinear problem: Numerical solution by the Newton-Raphson-Simpson method; need initial guess.

Initial guess: $p_{init} = [2, 600, 870]$ (2)

How ?

p_{init} = [2, 600, 870]

Result (using Octave):

$$p_{out} = [1.6407, 740.2473, 933.7323]$$

Least square curve was plotted against the data.

$$\mu_X = \mathbb{E}(X(\omega)) = \sum_{k=1}^{n} X(\omega_k) w(\omega_k)$$

\mu_X = {\mathbb E}(X (\omega))= \displaystyle \sum_{k=1}^{n} X(\omega_k) w(\omega_k)

Note: