Wiki Schubi

Summaries of all things computer science.

First Edition

PREFACE

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetuer.

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Part I

Fundamentals

CHAPTER 1

APPROXIMATION & NUMERICAL METHODS

1 Newton's Method

Definition 1.1: Iterative Newton Method

Newton's iterative method can be used for finding approximate roots of a function $f(x) = 0$. The iterative formula is given by:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \tag{1.1}$$

- x_n is the current approximation
- $f'(x_n)$ is the derivative of the function at x_n
- x_{n+1} is the next approximation

Definition 1.2: Multivariate Newton

$$(x) \approx f\left(x_0\right) + J_f\left(x_0\right) * \left(x - x_0\right) | x \approx x_0 \tag{1.2}$$

2 Taylor Series

Definition 2.1: Approximation Using Taylor Polynomial

Given a function f which is which is $N+1$ differentiable in range $[a;b]$ we can use Taylor polynomial

$$p_N(x) = \sum_{i=0}^{N} \frac{f^{(i)}(\hat{x})}{i!}(x-\hat{x})^i \tag{1.3}$$

around the middle of the development point

$$\hat{x} = \frac{a+b}{2} \tag{1.4}$$

to approximate. Further if

$$|f^{(N+1)}(x)| \leq m \tag{1.5}$$

Then the error can be assessed using:

$$|f(x) - p_N(x)| \leq \frac{m(b-a)^{N+1}}{2^{N+1}(N+1)!}, \forall x \in [a;b] \tag{1.6}$$

The series only accurately represents a function for those arguments x whose distance from the development point x_0 is smaller than the radius of convergence T of the Taylor series, i.e. for which the condition $|x - x_0| \leq R$ is fulfilled. The radius of convergence R is a distance from the function f and the development point x_0.

Definition 2.2: Linearization

$$y_t = f(x_0) + \left.\frac{\mathrm{d}f}{\mathrm{d}x}\right|_{x_0} \cdot (x - x_0), x \approx x_0 \tag{1.7}$$

This can be interpreted as a Taylor polynomial of degree one. This means we can choose a range I where the function f can be differentiated 2 times and choose m such that $\forall x \in I$:

$$|f^{''}(x)| \leq m \tag{1.8}$$

Then we can calculate the approximation error Δ:

$$x \in [x_0 - \sqrt{\frac{2\Delta}{m}}; x_0 + \sqrt{\frac{2\Delta}{m}}] \tag{1.9}$$

3 Fourier

Definition 3.1: Fourier Series

Let T be the fundamental period of a function $f(t)$. Then the quantity

$$\omega = \frac{2\pi}{T}$$

is called the fundamental angular frequency of this function (the term $k\omega$ is occasionally also referred to as ω_k).
If $f(t)$ is expressed in the form

$$f(t) = a_0 + \sum_{k=1}^{\infty} [a_k \cos(k\omega t) + b_k \sin(k\omega t)]$$

then this representation of f is called the Fourier series in the sine-cosine form. The coefficients a_0, a_k, and b_k are called the Fourier coefficients, and the collection of all Fourier coefficients is referred to as the sine-cosine spectrum of the Fourier series.
If $f(t)$ is expressed in the form

$$f(t) = a_0 + \sum_{k=1}^{\infty} A_k \cos(k\omega t + \varphi_k) \tag{1.10}$$

then this representation of the Fourier series is called the amplitude-phase form. In this case, the quantities A_k are called amplitudes, and φ_k are called phases of the Fourier series. The collection of all amplitudes is referred to as the amplitude spectrum, and the collection of all phases is referred to as the phase spectrum.

Definition 3.2: Coefficients

$$a_0 = \frac{1}{T}\int_{-T/2}^{T/2} f(t)\,dt$$

$$a_l = \frac{2}{T}\int_{-T/2}^{T/2} f(t)\cos(l\omega t)\,dt$$

$$b_l = \frac{2}{T}\int_{-T/2}^{T/2} f(t)\sin(l\omega t)\,dt$$

If $f(t)$ is an even function, then:

$$a_0 = \frac{2}{T}\int_{0}^{T/2} f(t)\,dt$$

$$a_l = \frac{4}{T}\int_{0}^{T/2} f(t)\cos(l\omega t)\,dt$$

$$b_l = 0$$

Thus, the Fourier series of an even function contains only the constant term and cosine terms, i.e., no sine terms.
If $f(t)$ is an odd function, then:

$$a_0 = 0$$

$$a_l = 0$$

$$b_l = \frac{4}{T}\int_{0}^{T/2} f(t)\sin(l\omega t)\,dt$$

Thus, the Fourier series of an odd function contains only sine terms, i.e., no constant term and no cosine terms.

Definition 3.3: Conversion

The conversion of Fourier coefficients from the sine-cosine form to the amplitude-phase form is performed using the following formulas:

$$A_k = \sqrt{a_k^2 + b_k^2},$$

$$\varphi_k = \begin{cases} \arccos\left(\frac{a_k}{A_k}\right) & \text{if } b_k \geq 0 \\ -\arccos\left(\frac{a_k}{A_k}\right) & \text{if } b_k < 0 \end{cases}$$

The phases φ_k have the following properties: - If $a_k > 0$ and $b_k = 0$, then $\varphi_k = 0$ - If $a_k < 0$ and $b_k = 0$, then $\varphi_k = \pi$ - If $b_k > 0$, then $\varphi_k \in (0, \pi]$ - If $b_k < 0$, then $\varphi_k \in [-\pi, 0)$.
The conversion from amplitude-phase form to sine-cosine form is performed using the following formulas:

$$a_k = A_k \cos(\varphi_k)$$

$$b_k = A_k \sin(\varphi_k)$$

Theorem 3.4: Dirichlet's Theorem

Let $f(t)$ be a piecewise continuously differentiable, T-periodic function, i.e., a T-periodic function that is continuously differentiable on the interval $[0, T]$ except at finitely many points, where the right-hand and left-hand limits exist at each discontinuity.
Then the Fourier series $S(t)$ of $f(t)$ agrees with the function $f(t)$ at all points where $f(t)$ is continuous. Specifically, at all points of continuity of $f(t)$, it holds that:

$$S(t) = f(t)$$

At points t where $f(t)$ has a discontinuity, the Fourier series converges to the average of the left-hand and right-hand limits. At such points, it holds that:

$$S(t) = \frac{1}{2}\left(\lim_{t \to t^+} f(t) + \lim_{t \to t^-} f(t)\right)$$

CHAPTER 2

WAHRSCHEINLICHKEIT UND STATISTIK

1 Kombinatorik

Die Kombinatorik ist eine Teildisziplin der Mathematik, die sich mit endlichen oder abzählbar unendlichen diskreten Strukturen beschäftigt und deshalb auch dem Oberbegriff Diskrete Mathematik zugerechnet wird.

Definition 1.1: Produktregel (Für-jede-gibt-es-Regel)

Wenn es für jede von n Wahlmöglichkeiten m mögliche weitere Wahlmöglichkeiten gibt, dann gibt es insgemsamt $n \cdot m$ Möglichkeiten.

$$n_1 \cdot n_2 \dots n_k = \prod_{i=1}^{k} n_i \tag{2.1}$$

Definition 1.2: Permutation (Anordnung)

Gegeben sind n anzuordnende Objekte. Die Fakultät ! beschreibt dabei die Anzahl Arten wie dies geschehen kann.

$$\begin{aligned} P_n &= \#\{\text{Plätze für 1. Objekt (n)}\} \cdot \#\{\text{Plätze für 2. Objekt (n-1)}\} \cdot \\ &\quad \dots \#\{\text{Plätze für n. Objekt (1)}\} \\ &= n \cdot (n-1) \dots 1 \\ &= n! \end{aligned} \tag{2.2}$$

Definition 1.3: Kombination (Auswahl)

Die Anzahl Objekte welche aus n von k möglichen ausgewählt werden (ohne zurücklegen) wird beschrieben durch:.

$$\begin{aligned} \frac{n!}{k!(n-k)!} &= \binom{n}{k} \\ \text{Binomialkoeffizient } (a+b)^n &= \sum^{n} \binom{n}{k} a^{n-k} b^k \\ \text{oder auch } (1+z)^n &= \sum^{n} \binom{n}{k} z^k \end{aligned} \tag{2.3}$$

Definition 1.4: Variation (Kerlenketten)

Die Arten auf welche man k mal unter n verschiedenen Objekten auswählen kann wird beschrieben durch:

$$= n \cdot n^{k-1} = n^k \tag{2.4}$$

1.1 Erzeugende Funktion

Ist die formale Potenzreihe einer Folge a_n. Diese kann verwendet werden um Zahlen a_n in Funktionen zu codieren. Damit lassen sich kombinatorische Probleme mit algebraischer Manipulation von erzeugenden Funktionen lösen. Z.B. Gegeben 3 Einfränkler, 4 Zweifränkler und 7 Fünfliber, auf wieviele Arten kann man 39 Franken bezahlen?

$$\begin{aligned}
\text{1Fr} &= z \\
\text{0-3 Einfränkler} &= f_1(z) = 1 + z + z^2 + z^3 \\
\text{0-4 Zweifränkler} &= f_2(z) = 1 + z^2 + z^4 + z^6 + z^8 \\
\text{0-7 Fünfliber} &= f_3(z) = 1 + z^5 + \cdots + z^{35} \\
\text{Alle } & f_1(z) \cdot f_2(z) \cdot f_5(z) = \\
&= 1 + z + 2z^2 + 2z^3 + \cdots + 4z^{23} + \cdots
\end{aligned} \tag{2.5}$$

2 Ereignisse

Definition 2.1: Elementarergeignis

Der Ausgang eines Experimentes heisst Elementarereignis, die Menge aller Elementarereignisse wird mit Ω bezeichnet. Ein Elementarereignis ω ist also ein Element von Ω, $\omega \in \Omega$.

Definition 2.2: Ereignis

Ist Ω eine Menge von Versuchsausgängen, dann heisst eine Teilmenge $A \subset \Omega$ ein Ereignis. Man sagt, das Ereignis A ist eingetreten, wenn bei einer Durchführung des Experimentes ein Versuchsausgang $\omega \in A$ aufgetreten ist.

Definition 2.3: Ereignisalgebra

Eine Ereignisalgebra (Ω, A) ist eine Menge Ω mit einer Menge $A \subset Potenzmenge(\Omega)$ von Teilmengen von Ω , die folgende Bedingungen erfüllen:

1. Vereinigungen von Elementen von A sind ebenfalls in A, also
$$B, C \in A \Rightarrow B \cup C \in A$$
2. Differenzen von Elementen von A sind in A, also
$$B, C \in A \Rightarrow B \setminus C \in A$$
3. $\Omega \in A$ d.h. es gibt das sichere Ereignis.

Sind nur die Bedingungen 1 und 2 erfüllt, spricht man auch von einem Mengen-Ring. Eine Ereignisalgebra heisst manchmal auch ein Mengenkörper. Aus den Axiomen folgt direkt:

1. Es gibt das unmögliche Ereignis: $\emptyset$
2. Das Komplement eines Ereignisses ist ebenfalls ein Ereignis
3. Der Durchschnitt zweier Ereignisse ist ebenfalls ein Ereignis.

$$\begin{aligned} A \cap (B \cup C) &= (A \cap B) \cup (A \cap C) \\ A \cup (B \cap C) &= (A \cup B) \cap (A \cup C) \\ \overline{A \cap B} &= \bar{A} \cup \bar{B} \text{ and vice versa} \end{aligned} \tag{2.6}$$

3 Wahrscheinlichkeit

Die Wahrscheinlichkeit soll ein Mass dafür sein, dass ein Ereignis eintritt. Es gibt verschiedene Ansätze, wie wir zu einem solchen Mass kommen könnten:

1. Je häufiger ein Ereignis eintritt, desto grösser sollte die Wahrscheinlichkeit sein. Dies setzt voraus, dass das Experiment im Prinzip beliebig oft wiederholbar ist. Man nennt dies den frequentistischen Ansatz.

2. Die Wahrscheinlichkeit ist ein Mass für die persönliche Überzeugung, dass ein Ereignis eintreten wird. Im Unterschied zum frequentistischen Ansatz sollte man bei diesem Ansatz einem Ereignis auch dann eine Wahrscheinlichkeit geben können, wenn sich ein Experiment nicht wirklich wiederholen lässt. Dies ist der Bayessche Ansatz.

3. Wir könnten eine Reihe von plausiblen Axiomen postulieren, nach denen sich die Wahr scheinlichkeit zu verhalten hat, und dann zu untersuchen, ob ein solches Objekt tatsäch lich existiert. Dabei ist uns egal, was der Wahrscheinlichkeitswert genau bedeutet. Dieser axiomatische Ansatz hat den Vorteil, logisch konsistent zu sein, was bei den anderen An sätzen nicht von vornherein garantiert ist.

In allen Fällen ergibt sich eine Reihe von Gesetzmässigkeiten oder Formeln, welche die jeweiligen Wahrscheinlichkeitsbegriffe erfüllen müssen. Soll die Wahrscheinlichkeit eine objektive Grösse sein, dann muss für wiederholbare Experimente die für den Bayesschen Ansatz nötige persönliche Überzeugung direkt mit der Häufigkeit des Eintretens zusammenhängen. Der frequentistische und der Bayessche Ansatz werden also in diesem Fall übereinstimmen. Die Axiome im axiomatischen Ansatz sind natürlich genau die Rechenregeln, die man sowohl im frequentistischen Ansatz wie auch im Bayesschen Ansatz von der Wahrscheinlichkeit erwartet. Man darf daher davon ausgehen, dass die drei Ansätze die gleichen numerischen Resultate lie fern, sie unterscheiden sich höchstens in der Interpretation der Resultate.

Definition 3.1: Wahrscheinlichkeit als relative Häufigkeit

$$P(A) = \lim_{N\to\infty} \frac{n}{N} \tag{2.7}$$

Definition 3.2: Axiome eines Wahrscheinlichkeitsraumes

Wertebereich

$$0 \leq P(A) \leq 1 \tag{2.8}$$

Sicheres Ereignis

$$P(\Omega) = 1 \tag{2.9}$$

Vereinigung. Sind die Ereignisse $A_1, A_2, \dots$ paarweise disjunkt. also $A_i \cap A_j = \emptyset$ für $i \neq j$, dann gilt

$$P(A_1 \cup A2 \cup \dots) = P(A_1) + P(A_2) + \dots \tag{2.10}$$

Aus den Axiomen folgt:

1. Die Wahrscheinlichkeit des unmöglichen Ereignisses ist

$$P(\emptyset) = 0$$

2. Die Wahrscheinlichkeit des komplementären Ereignisses ist

$$P(\bar{A} = P(\Omega \setminus A) = 1 - P(A)$$

3. Die Wahrscheinlichkeit der Differenz der Ereignisse A und B ist

$$P(A \setminus B) = P(A) - P(A \cap B)$$

4. Die Wahrscheinlichkeit der Vereinigung zweier beliebiger Ereignisse ist (Ein-/Ausschaltformel)

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Definition 3.3: Laplace Experiment

Ein Experiment mit $n = |\Omega|$ Ausgängen heisst ein Laplace-Experiment, wenn jeder Versuchsausgang gleich wahrscheinlich mit Wahrscheinlichkeit

$$P(\omega) = \frac{1}{n}, w \in \Omega \tag{2.11}$$

$$P(A) = \frac{|A|}{|\Omega|} \tag{2.12}$$

Nur das unmögliche Ereignis hat Wahrscheinlichkeit 0.

Definition 3.4: Bernoulli Experiment

Genau zwei Versuchsausgänge mit Wahrscheinlichkeiten p und $1 - p$.

Definition 3.5: Bedingte Wahrscheinlichkeit

Die bedingte Wahrscheinlichkeit eines Ereignisses A unter der Bedingung B ist

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{2.13}$$

Man liest dies auch als “Wahrscheinlichkeit von A bedingt B”.

Definition 3.6: Unabhängigkeit

Die Ereignisse A und B heissen unabhängig, wenn gilt:

$$P(A \cap B) = P(A) \cdot P(B) \tag{2.14}$$

Definition 3.7: Satz der totalen Wahrscheinlichkeit

Ist B_i eine Folge paarweise disjunkter Mengen mit $\bigcup_{i=0}^{n} B_i = \Omega$, dann gilt für jedes Ereignis A

$$P(A) = \sum_{i=0}^{n} P(A|B_i) \cdot P(B_i) \tag{2.15}$$

Definition 3.8: Satz von Bayes

Für zwei beliebige Ereignisse mit A und B mit nicht verschwindender Wahrscheinlichkeit $P(B) \neq 0$ gilt

$$\begin{aligned} P(A|B) \cdot P(B) = P(A \cap B) &= P(B|A) \cdot P(A), \\ P(A|B) &= \frac{P(B|A) \cdot P(A)}{P(B)} \end{aligned} \tag{2.16}$$

4 Zufallsvariable

In der Stochastik ist eine **Zufallsvariable** (auch **zufällige Variable**, **zufällige Größe**, **zufällige Veränderliche**, **zufälliges Element**, **Zufallselement**, **Zufallsveränderliche**) eine Größe, deren Wert vom Zufall abhängig ist. Formal ist eine Zufallsvariable eine Funktion, die jedem möglichen Ergebnis eines Zufallsexperiments eine Größe zuordnet. Ist diese Größe eine reelle Zahl, so spricht man von einer **reellen Zufallsvariablen** oder **Zufallsgröße**. Beispiele für reelle Zufallsvariablen sind die Augensumme von zwei geworfenen Würfeln und die Gewinnhöhe in einem Glücksspiel. Zufallsvariablen können aber auch komplexere mathematische Objekte sein, wie Zufallsfelder, Zufallsbewegungen, Zufallspermutationen oder Zufallsgraphen. Über verschiedene Zuordnungsvorschriften können einem Zufallsexperiment auch verschiedene Zufallsvariablen zugeordnet werden.

Definition 4.1: Zufallsvariable

Eine Zufallsvariable ist eine Funktion $Q \rightarrow \mathbb{R}$. Eine diskrete Zufallsvariable nimmt nurdiskrete Werte in $\mathbb{R}$ an. Bei einer stetigen Zufallsvariable sind beliebige Werte $X(\omega) \in (R)$ möglich. Mit der zusätzlichen Eigenschaft , dass die Mengen $\{\omega | X(\omega) < x\}$ Ereignisse sind.

Mit einer Zufallsvariablen X kann man wieder neue Ereignisse definieren.

$$\begin{aligned} \{X = a\} &= \{\omega \in \Omega | X(\omega) = a\} \\ \{a < X \leq b\} &= \{\omega \in \Omega | a < X(\omega) \leq b\} \end{aligned} \tag{2.17}$$

Bei einer stetigen Zufallsvariable sind die Ereignisse der Form $\{X = a\}$ nur von beschränktem Nutzen, da nur ganz wenige Versuche dazu führen werden, dass das Ereignis eintritt.

5 Erwartungswert

Der **Erwartungswert** (selten und doppeldeutig *Mittelwert*) ist ein Grundbegriff der Stochastik. Der Erwartungswert ist eine Kennzahl einer Zufallsvariablen. Bei einer engeren Definition ist der Erwartungswert einer Zufallsvariablen eine reelle Zahl und damit endlich; bei einer weiteren Definition sind für den Erwartungswert einer Zufallsvariablen auch die Werte $\pm\infty$ zugelassen. Es gibt Zufallsvariablen, für die kein Erwartungswert definiert ist.

Definition 5.1: Erwartungswert

Sei X eine Funktion auf Ω , und lasse sich Ω in endlich viele Ereignisse A_i zerlegen, auf denen $X(\omega)$ konstant ist, dann ist der Erwartungswert von X

$$E(X) = \sum_{i=0}^{n} P(A_i) \cdot X(A_i) \tag{2.18}$$

Der Erwartungswert einer reelen Zufallsvariable ist diejenige Uahl, für die die Varianz minimal wird.
Sind X und Y Zufallsvariable mit Werten in $\mathbb{R}$, und λ, dann gilt

1. $E(X+Y) = E(X) + E(Y)$
2. $E(\lambda X) = \lambda E(X)$
3. Sei X_A die charakteristische Funktion des Ereignisses $A \in \mathbb{A}$, welche definiert ist durch

$$X_A : \Omega \to \mathbb{R} : \omega \mapsto \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases}$$

dann gilt $E(X_A) = P(A)$.

Definition 5.2: Unabhängigkeit

Zwei Zufallsvariable X und Y heissen unabhängig, wenn die Ereignisse $\{\omega|X(\omega) \leq x\}$ und $\{\omega|Y(\omega) \leq y\}$ für alle $x, y \in \mathbb{R}$ unabhängig sind, also

$$P((X \leq x) \wedge (Y \leq y)) = P(X \leq x)P(Y \leq y) \quad \forall x, y \in \mathbb{R} \tag{2.19}$$

Definition 5.3: Erwartungswert Produkregel

Sind X und Y unabhängige Zufallsvariablen, dann gilt

$$E(XY) = E(X) \cdot E(Y)$$

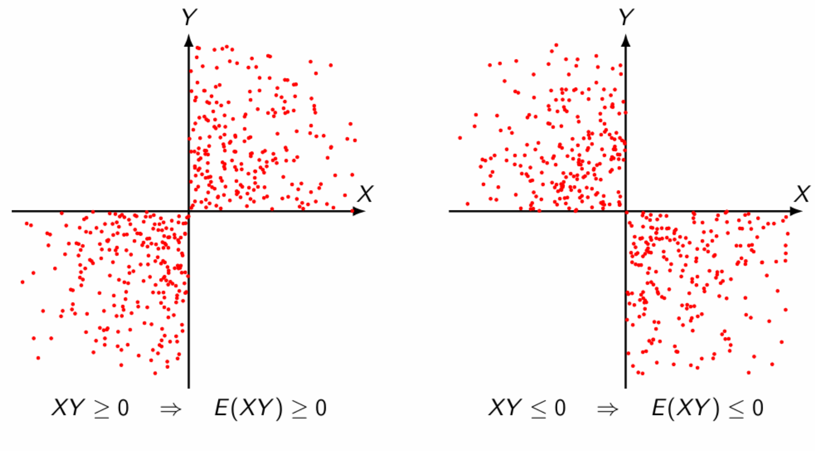


Figure 2.1: Unabhängigkeit

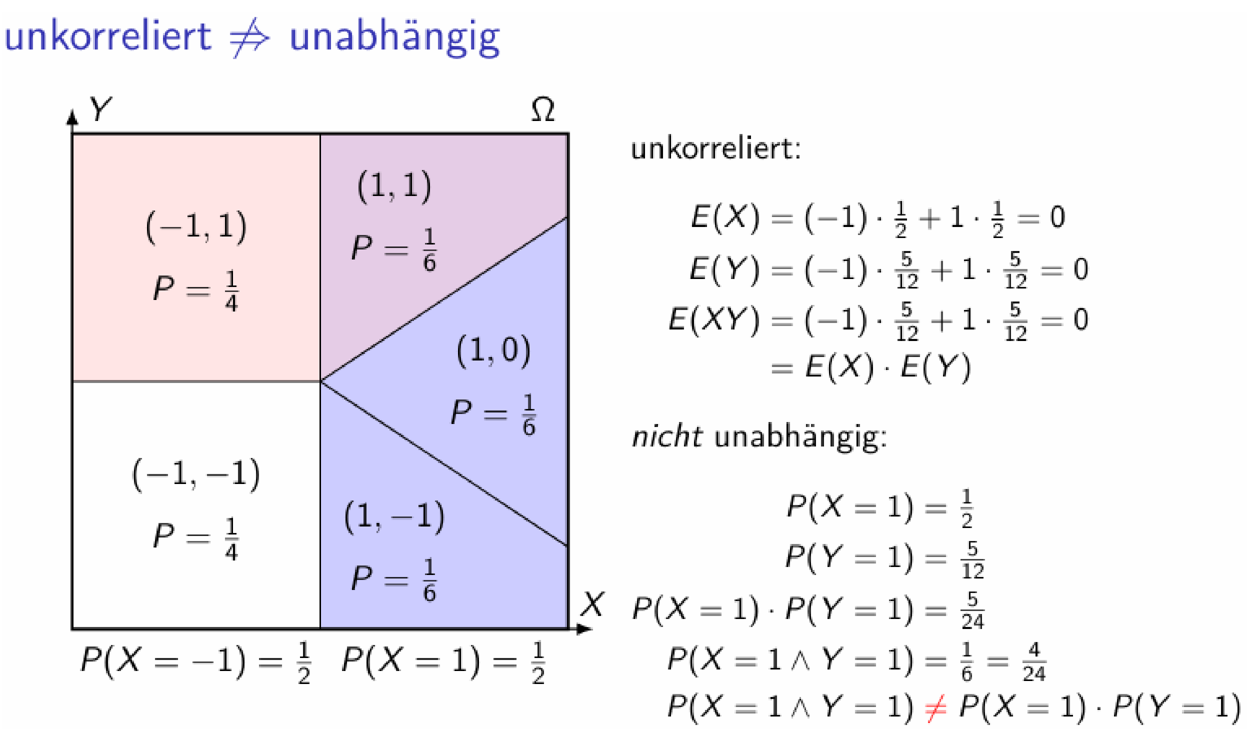


Figure 2.2: Unkorreliert - Unabhängig

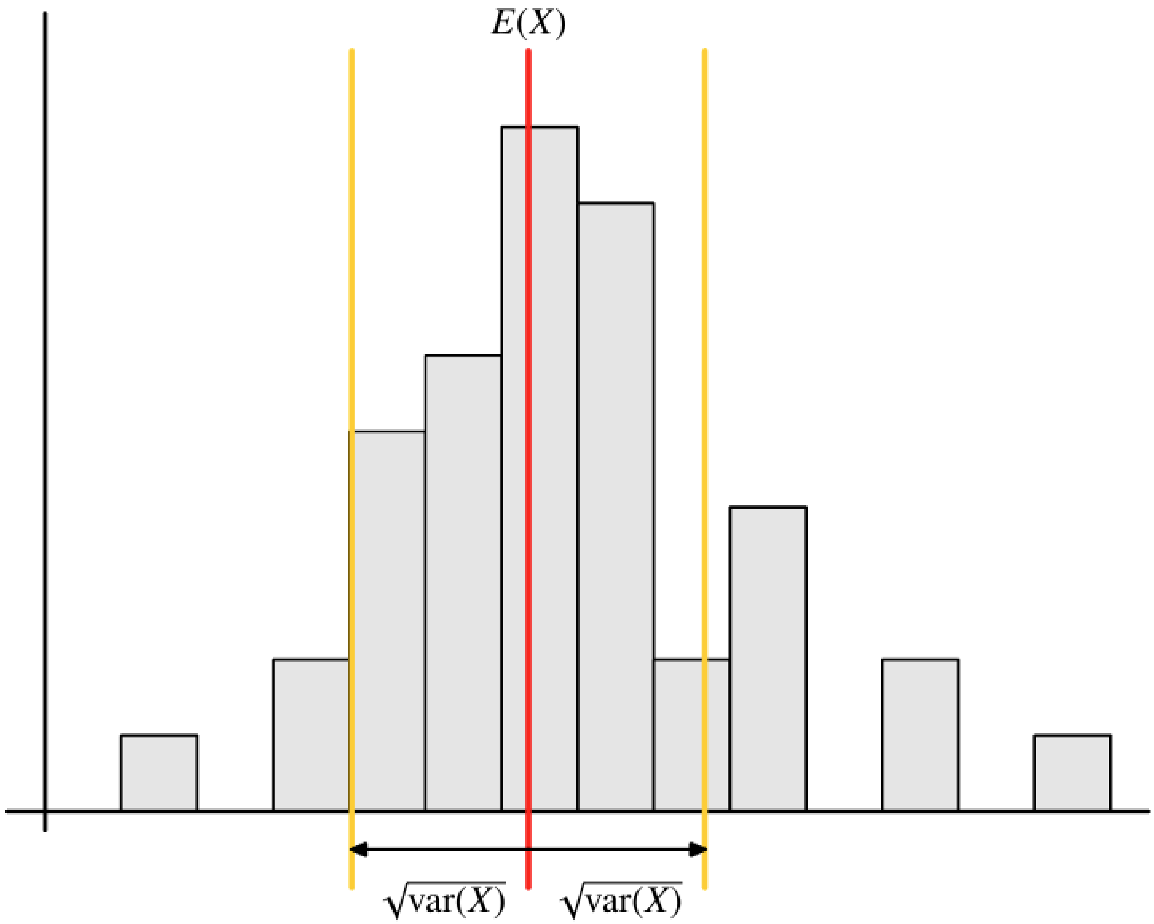


Figure 2.3: Darstellung von Erwartungswert und Varianz

6 Varianz

Die **Varianz** (lateinisch *variantia* „Verschiedenheit“ bzw. *variare* „(ver)ändern, verschieden sein“) ist ein Maß für die Streuung einer Wahrscheinlichkeitsdichte um ihren Schwerpunkt. Mathematisch wird sie definiert als die mittlere quadratische Abweichung einer reellen Zufallsvariablen von ihrem Erwartungswert. Sie ist das zentrale Moment zweiter Ordnung einer Zufallsvariablen.

Definition 6.1: Varianz

Sei $X : \Omega \to \mathbb{R}$ eine Zufallsvariable, dann heisst die durch $var(X)^2 = E((X - E(X))^2)$ definierte Grösse var(X) die Varianz von X. Oder auch die mittlere quadratische Abweichung vom Erwartungswert. Es ist insbesondere

$$var(X) = E((X - E(X))^2) = E(X^2) - E(X)^2 \tag{2.20}$$

Man beachte, dass $var(X)$ die Masseinheit des Quadrates von X hat. Wenn also X als Masseinheit eine Länge hat, dann hat $var(X)$ die Masseinheit einer Fläche. Insbesondere kann man $var(X)$ nicht in der gleichen Zeichnung visualisieren wie $E(X)$. Aber die Grösse $\sqrt{var(X)}$ hat die gleiche Masseinheit, sie drückt die “Streubreite” der Werte von X.

Definition 6.2: Rechenregeln

Seien X und Y unabhängige Zufallsvariable, dann haben Summe und Produkt folgende Varianz

$$\begin{aligned} var(\lambda X) &= \lambda^2 var(X) \\ var(X + Y) &= var(X) + var(Y) \\ var(XY) &= var(X)var(Y) + var(Y)E(X)^2 + var(X)E(Y)^2 \end{aligned} \tag{2.21}$$

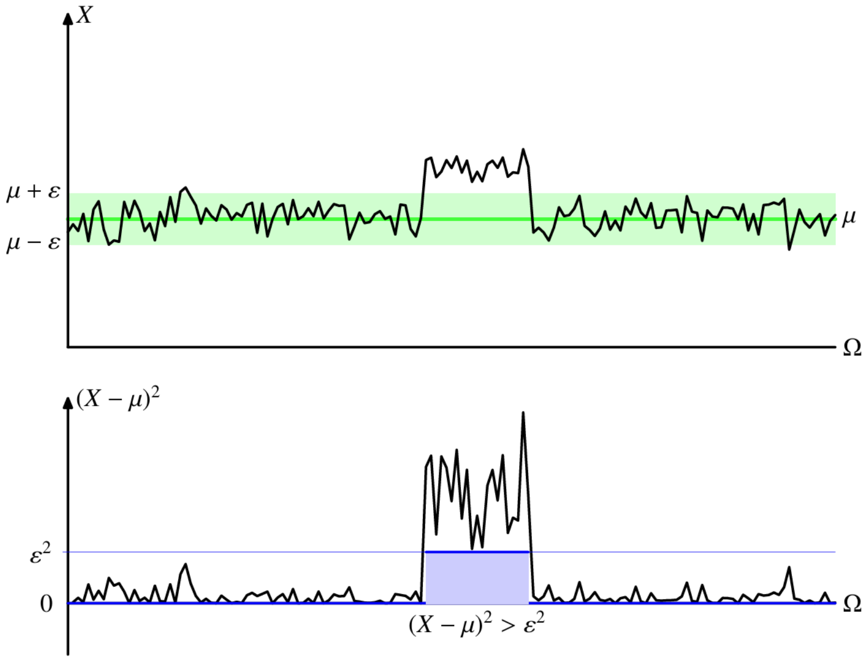


Figure 2.4: Herleitung der Tschebyscheff-Ungleichung, charakteristische Funktion des Ereignisses |X - µ| > ϵ ist blau eingezeichnet.

7 Bewertung vom Mittelwert / Erwartungswert

Definition 7.1: Tschebyscheff Ungleichung

X eine Zufallsvariable mit Erwartungswert $\mu = E(X)$, dann lässt sich die Wahrscheinlichkeit, dass X um mehr als ϵ vom Erwartungswert abweicht, wie folgt abschätzen:

$$P(|X - \mu| > \varepsilon) \leq \frac{var(x)}{\varepsilon^2} \tag{2.22}$$

Lässt man ε wachsen findet man, dass Abweichungen deutlich grösser als $\sqrt{var(X)}$ unwahrscheinlich sind. Leider sind die Vorhersagen wie auch bei der Tschebyscheff-Ungleichung nur beschränkt nützlich. Wenn die Wahrscheinlichkeit, dass der Mittelwert um mehr als σ vom Erwartungswert abweicht, kleiner als 1% sein soll, sind dafür nach dem Gesetz der grossen Zahlen mindestens 100 Summanden nötig.

Definition 7.2: Bernoulli Gesetz der grossen Zahlen

Die Wahrscheinlichkeit, dass der Mittelwert von n unabhängigen Zufallsvariablen mit Erwartungswert μ und Varianz σ^2 mehr als ε von μ abweicht, ist :

$$P(|M_n - \mu| > \varepsilon) \leq \frac{var(X)}{n\varepsilon^2} \tag{2.23}$$

Insbesondere gilt

$$\lim_{n\to\infty} P(|M_n - \mu| > \varepsilon) = 0 \tag{2.24}$$

Definition 7.3: Genauigkeit der empirischen Häufigkeit

Wird ein Experiment n mal durchgeführt, und tritt dabei das Ereignis A mit der relativen Häufigkeit h ein, dann ist die Wahrscheinlichkeit, dass h um mehr als ε von $P(A)$ abweicht

$$P(|h - P(A)| > \varepsilon) \leq \frac{P(A)(1 - P(A))}{n\varepsilon^2} \leq \frac{1}{4n\varepsilon^2} \tag{2.25}$$

8 Lineare Regression

9 Verteil- und Dichtefunktionen

Definition 9.1: Verteilfunktion

Die Verteilungsfunktion beschreibt die Wahrscheinlichkeiten der Werte einer Zufallsvariable:

$$F(x) = P(X \leq x) \tag{2.26}$$

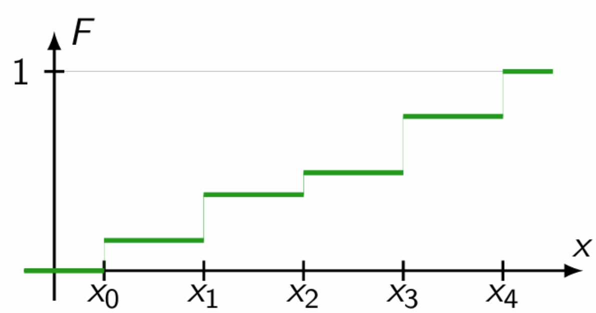


Figure 2.5: Diskrete Verteilfunktion

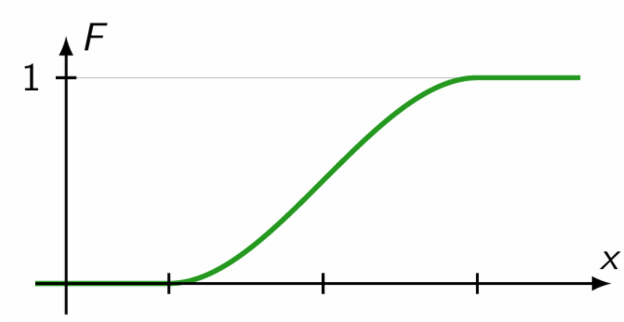


Figure 2.6: Stetige Verteilfunktion

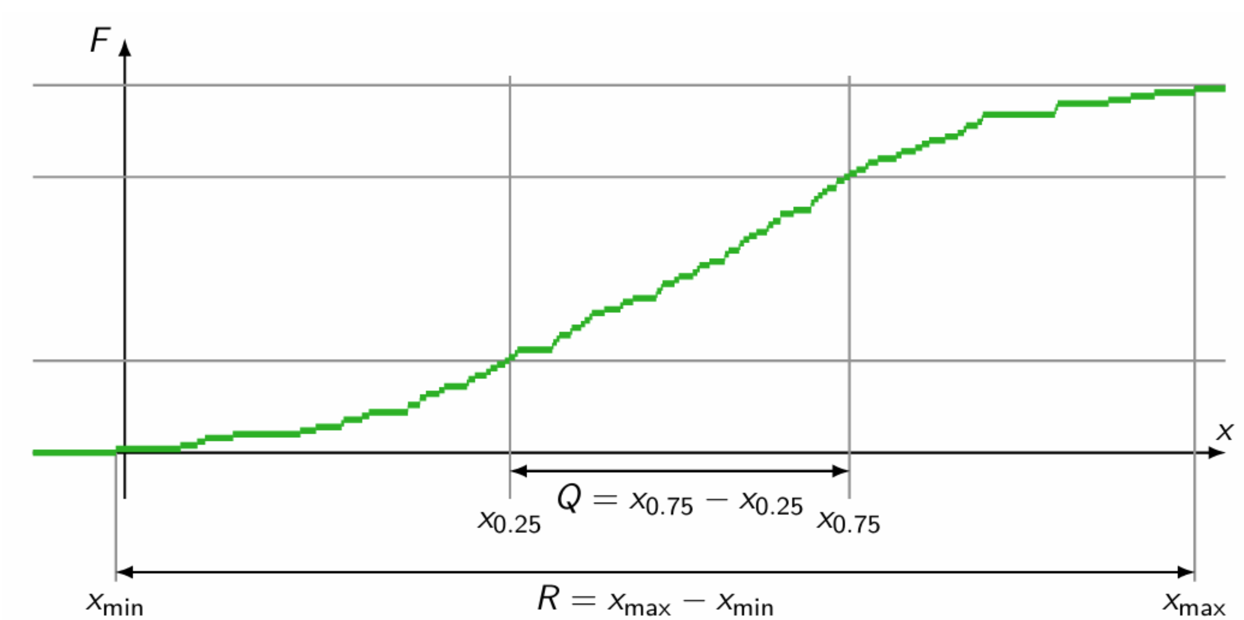


Figure 2.7: Spannweite R und Quartilabstand Q

9.1 Wahrscheinlichkeitsdichte

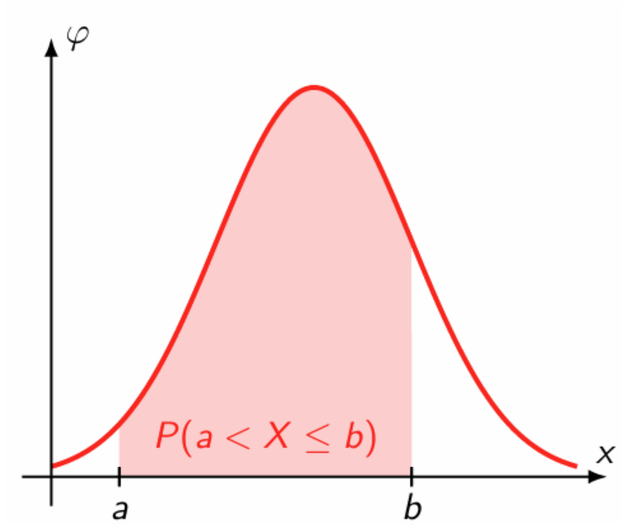


Figure 2.8: Wahrscheinlichkeitsdichte

Definition 9.2: Wahrscheinlichkeitsdichte Funktion

Die Ableitung der Verteilungsfunktion heisst Wahrscheinlichkeitsdichte:

$$\varphi(x) = \frac{d}{dx}F(x) = F'(X) \tag{2.27}$$

Wenn $F(x)$ (Verteilfunktion) die Stammfunktion ist bedeutet dies:

$$P(a < X \leq b) = \int_a^b \varphi(x)dx \tag{2.28}$$

Mit $\varphi(x)$ kann man nun auch $E(X)$ berechnen:

$$\begin{aligned} E(X) &= \sum x \cdot p(x) \\ &= \int_{-\infty}^{\infty} x \cdot \varphi(x)dx \end{aligned} \tag{2.29}$$

9.2 Gleichverteilung kreieren

Durch erneute Anwendung der Verteilfunktion auf X wird die Zufallsvariable gleichverteilt.

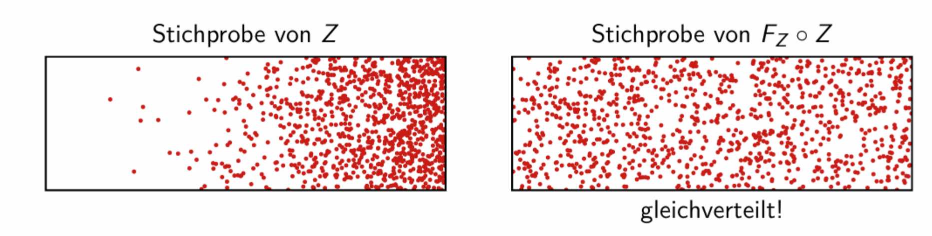


Figure 2.9: Enter Caption

10 Gleichverteilung

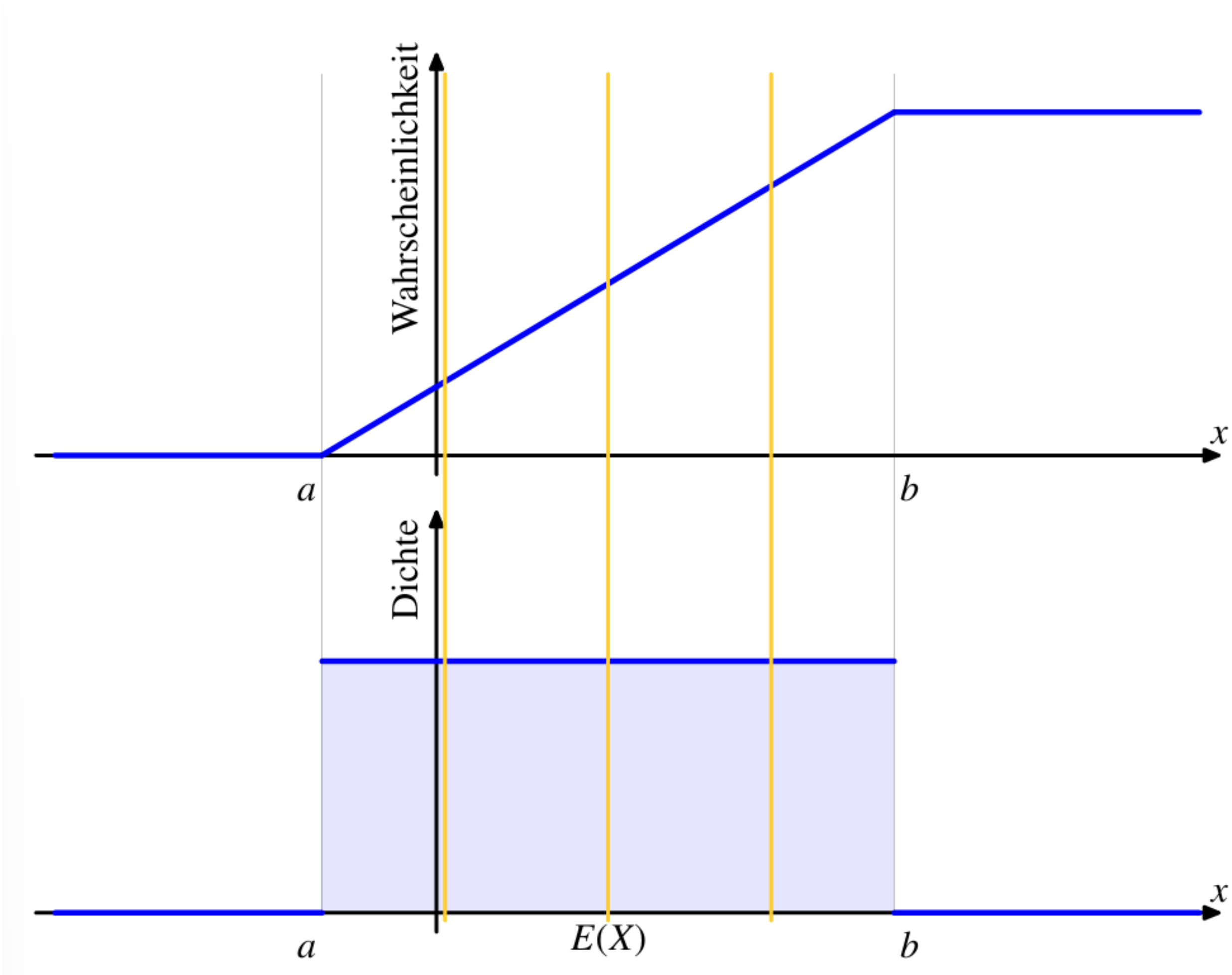


Figure 2.10: Gleichverteilung

Definition 10.1: Dichtefunktion

$$\varphi(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{sonst} \end{cases} \tag{2.30}$$

Definition 10.2: Verteilfunktion

$$F(x) = \begin{cases} 0 & x \leq a \\ \frac{x-a}{b-a} & x \leq a \leq b \\ 1 & x > b \end{cases} \tag{2.31}$$

Definition 10.3: Erwartungswert

$$E(X) = \frac{a+b}{2} \tag{2.32}$$

Definition 10.4: Varianz

$$\sigma^2 = \frac{(b-a)^2}{12} \tag{2.33}$$

Definition 10.5: Median

$$x_{\frac{1}{2}} = \frac{a+b}{2} \tag{2.34}$$

Definition 10.6: Wahrscheinlichkeit für Abweichung

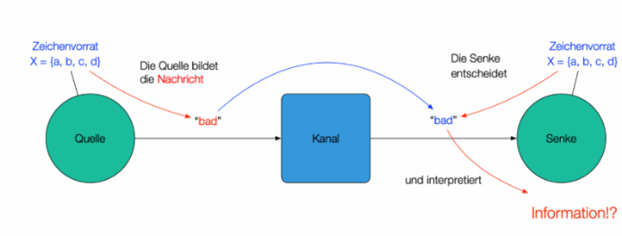
$$P(|X - E(X)| > \epsilon) = 1\frac{2\epsilon}{b-a} \text{ für } \epsilon < \frac{b-a}{2} \tag{2.35}$$

Part II

Theory of Codes, Information and Computation

CHAPTER 1

INFORMATION THEORY / CODING



Nachricht (Darstellung & Bedeutung)	redundant	nicht-redundant
irrelevant	Zeichenvorrat bei Quelle und Senke verschieden	
relevant	vorhersagbar	Information

Figure 1.1: Information Theory Term Overview

1 Information Theory

Definition 1.1: Decision Content / Entscheidungsgehalt

The measure of the effort required to form a message or to decide on a decision of a message is the decision content.

$$H_0 = log_2(N)[\text{bit}] \tag{1.1}$$

Definition 1.2: Decision Flow / Entscheidungsfluss

The formula measures the rate of information flow necessary to make a decision over time.

$$H_0 = \frac{log_2(N)}{\tau} \tag{1.2}$$

Where τ is the time required for the transmission of a word.

Definition 1.3: Information Content

The information content of a sign indicates how many elementary decisions have to be made to determine this sign.

$$I(x_k) = log_2(\frac{1}{P(x_k)}) \tag{1.3}$$

Definition 1.4: Entropy (mittlerer Informationsgehalt)

Entropy is the average information content of the source. It therefore shows how many elementary decisions the source/sink has to make on average per character

$$H(X) = \sum_{k=1}^{N} P(x_k) \cdot I(x_k) = \sum_{k=1}^{N} p(x_k) \cdot log_2(\frac{1}{P(x_k)}) \tag{1.4}$$

Since we typically use a binary encoding for the discrete values of the source we use:

$$L(X) = \sum_{k=1}^{N} P(x_k) \cdot I_r(x_k) \tag{1.5}$$

where $I_r = I(x_k)$ aufgerundet (auch tatsächliche Codewortlänge)

Definition 1.5: Source Redundancy (Redundanz der Quelle)

$$R_Q = H_0 - H(X) \tag{1.6}$$

Definition 1.6: Prefix Code

A prefix code is a code that fulfills the Fano condition: No code word of the code is a prefix of another code word. In other words, no code word may be the beginning of another code word. For example, a code with the code words {0, 10, 11} fulfills the prefix property, whereas the code with the code words {0, 01, 10} does not, as 0 is the prefix of 01.

Theorem 1.7: Shannons Theorem

Für jede beliebige zugehörige Binärcodierung mit Präfixeigenschaft ist die mittlere Codewortlänge nicht kleiner als die Entropie $H(X)$.

$$H(x) \leq L \tag{1.7}$$

Für jede beliebige Quelle kann eine Binärcodierung gefunden werden, so dass die folgende Ungleichung gilt:

$$H(X) \leq X \leq H(X) + 1 \tag{1.8}$$

Definition 1.8: Code Redundancy

$$R_C = L - H(X) \tag{1.9}$$

Definition 1.9: Discrete Source with and without memory

Die mittlere Entropie einer Quelle ohne Gedächtnis ist stets grösser oder gleich der Entropie einer Quelle mit Gedächtnis. In der Quellencodierung sind daher nicht Einzelzeichen zu codieren, sondern stets Zeichenketten.

$$R_Q = H_0 - H_{oG}(X) \leq H_= - H_{mG}(X) \tag{1.10}$$

2 Huffman

Verfahren zur Entwicklung eines kommafreien Codes mit minimaler mittlerer Codewortlänge Rekursives Verfahren, d.h. der Binärbaum wird nicht von der Wurzel, sondern von den Blättern aus entwickelt

3 Lempel-Ziv

Example : Algorithm

1. Suche in der Tabelle eine möglichst lange Zeichenfolge, die mit den nächsten N zu kodierenden Zeichen übereinstimmt.
2. Bilde ein Token und speichere es.
3. Verschiebe das Fenster um $N + 1$ Zeichen.
4. Wiederhole, bis alle Zeichen kodiert sind.

4 Kanalmodell

TODO

Part III

Algorithms and Datastructures (AD)

Part IV

Data Science

CHAPTER 1

STATISTICAL MACHINE LEARNING

1 Basics

Definition 1.1: Correlation and Causation

REMEMBER! Correlation does not imply causation!

1.1 Prediction vs Inference

Definition 1.2: Statistical Machine Learning

The main goal of statistical learning is to minimize the reducible error.

Definition 1.3: Error

Output (dependable variable) depends not only on the input, but also a from X independend zero meanr error term ϵ. Even a perfect prediction function $f(x_i)$ has a **irreducible error** ϵ.

- Random Measurement Error
- Systematic Errors e.g. humidity

If the estimated $\hat{f}(X)$ is not equal to the true function, there is the possibility to learn a better relationship between input and output, this is called **reducible error**.

$$\begin{aligned} MSE &= E\{(Y-\hat{Y})^2\} \\ &= (1)E\{(f(X)-\hat{f}(X))^2\} + (2)Var(\epsilon) \\ (1)\text{Reducible Error } &+ (2)\text{Irreducible Error} \end{aligned} \tag{1.1}$$

Definition 1.4: Inference

In inference we are interested in understanding the relationship between predictors and response.

Definition 1.5: Estimation Aproaches

Parametric Assumptions are made about the functional form of $f(X)$, where some missing parameters need to be learned. For example a simple linear model assumes form: $f(X) = \beta_0 + \beta_1 X_1$. If the form of $\hat{f}(X)$ is far from $f(X)$, then this approach achieves badly.

Non-Parametric Methods that make no assumptions about the functional form of $f(X)$.

Definition 1.6: Supervised Learning

In supervised learning for each input vector X, its corresponding dependent variable Y is known.

2 Bias-Variance

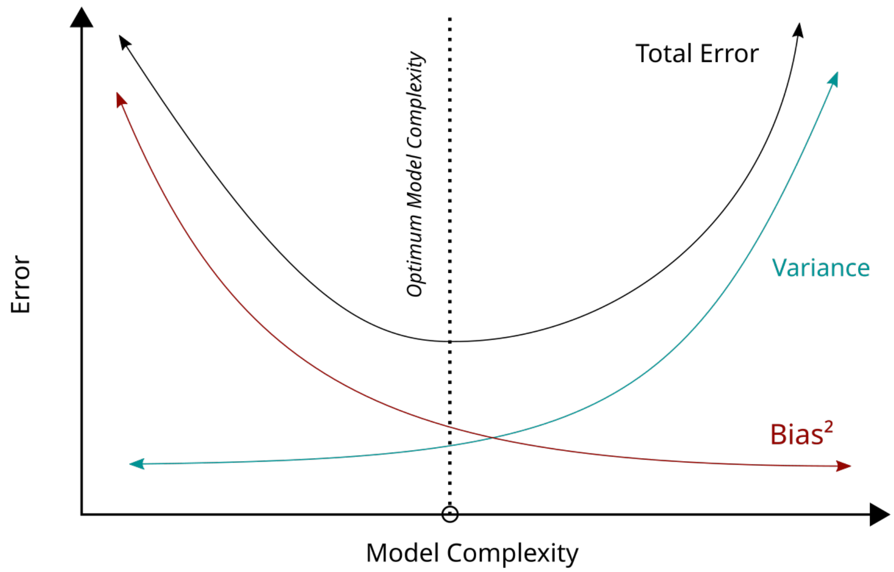


Figure 1.1: Bias Variance Tradeoff

The expected test MSE is the sum of a variance term, a bias term squared and the irreducible error. Its defined as the average test MSE that would be obtained if $f(X)$ is estimated repeatedly using a large number of training sets at tested at x_0.

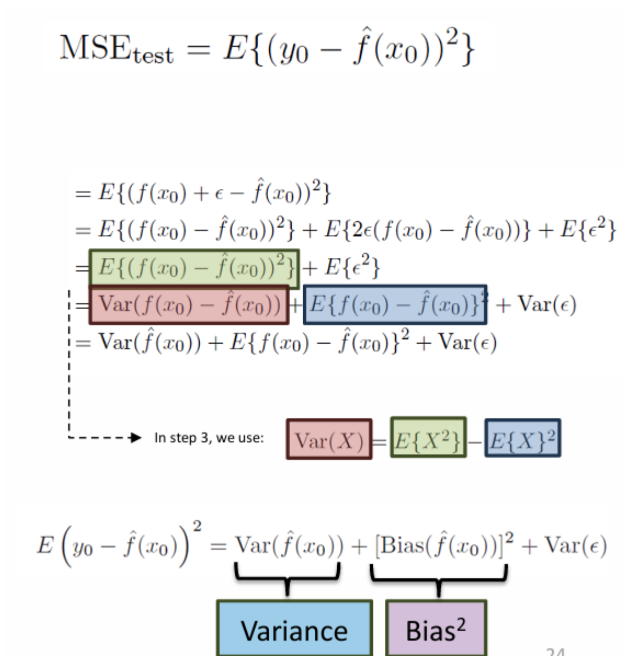


Figure 1.2: Enter Caption

The $\sqrt{variance}$ refers to the amount by which the estimate would change, if we estimated using a different training set. Ideally the training set should have little influence on how the method estimates, this is typical for low flexibility methods.

The bias refers to the error introduces by approximating. If a model is too simple (low flexibility) to capture the true shape of the function, this will result in high bias.

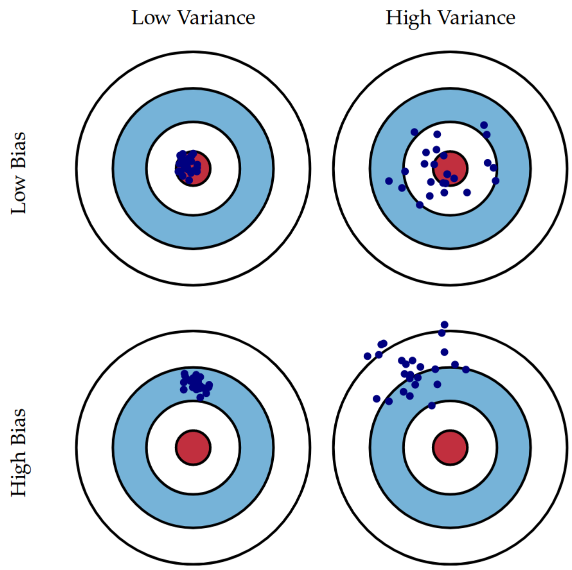


Fig. 1 Graphical illustration of bias and variance.

Figure 1.3: Bias Variance Tradeoff

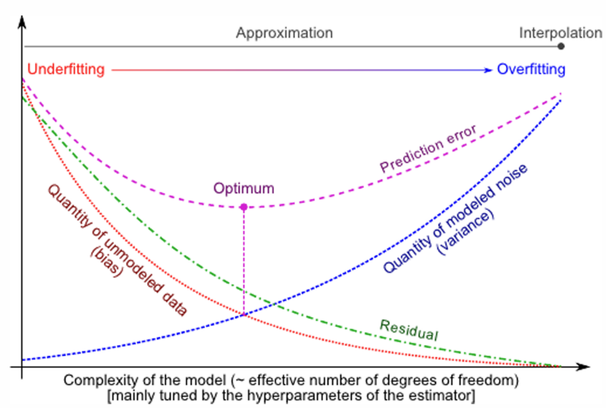


Figure 1.4: Bias Variance Tradeoff Optimum

Definition 2.1: Bias-Variance Summary

- Flexible methods have low bias but high variance
- Simple methods have low variance but high bias

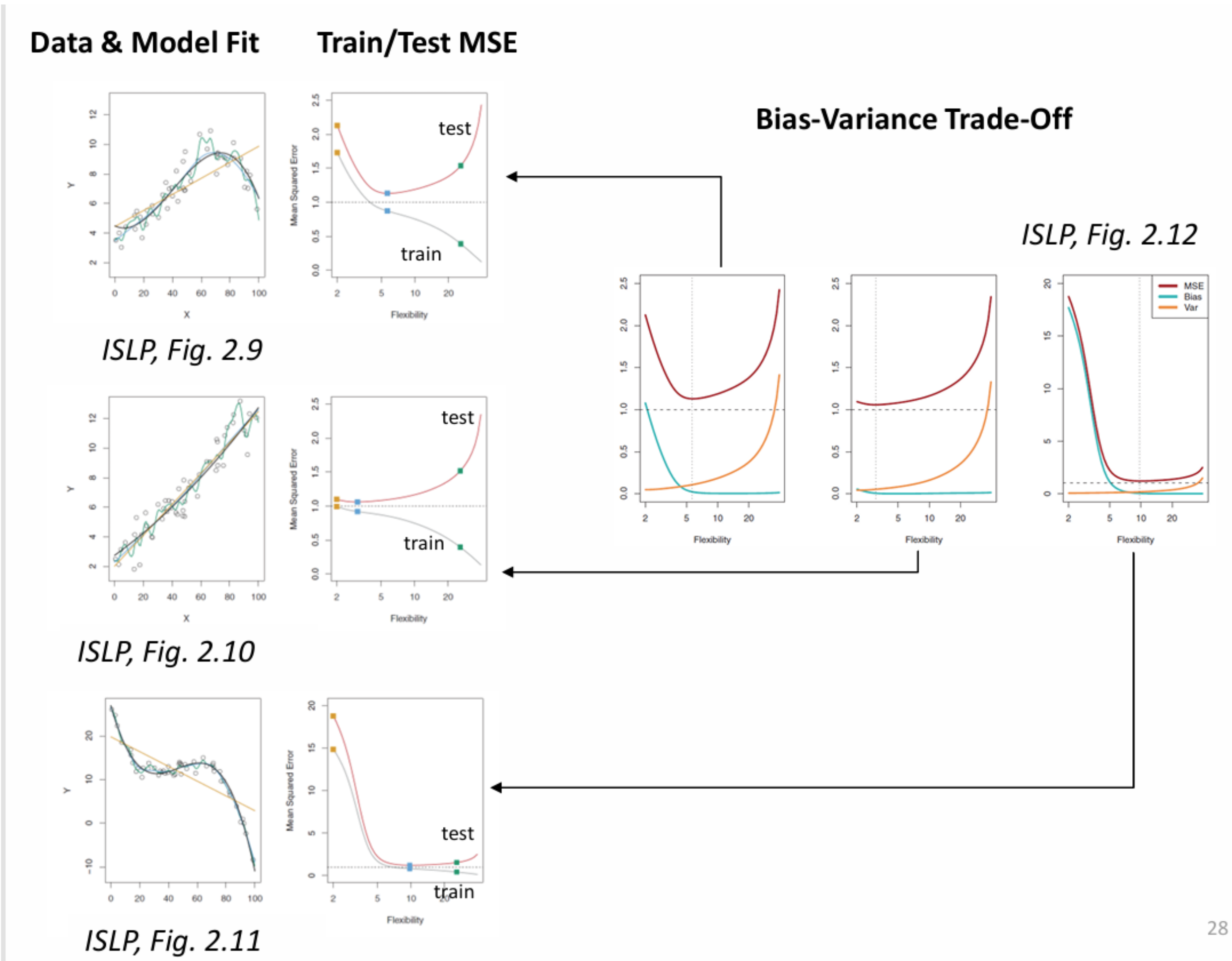


Figure 1.5: Bias Variance Overview

3 Regression

Quantitative (numerical) prediction where Y is an ordered numerical value.

Definition 3.1: Mean Squared Error

$$MSE = \frac{1}{N}(y_i - \hat{f}(x_i))^2 \tag{1.2}$$

4 Classification

Qualitative (categorical) prediction where Y is discrete categorical value.

Definition 4.1: Classification Error Rate

$$Error = \frac{1}{n}\sum_{i=1}^{n} I(y_i \neq \hat{y}_i) \tag{1.3}$$

Definition 4.2: Bayes Classifier

Test error rate is minimized, by a classifier that assigns each observation to the most probable class, given its predictor value. Results in the lowest possible error rate called Bayes error rate.

$$\begin{gathered} Pr(Y = j|X = x_0) \\ \text{set class to: } max_j(Pr(Y = j|X = x_0)) \end{gathered} \tag{1.4}$$

Where j is the most probable class for a given input variable (predictor) $X = x_0$. E.g for a 2 class model: $Pr(Y = 1|X = x_0) > 0.5$
The bayes error rate is analogous to the irreducible error and is 0.1304.

Definition 4.3: K-Nearest Neighbors

Bayes classifier requires knowledge of the conditional distribution of $Y|X$, in a real world problem this is never known. K-nearest neighbors is the simplest of such methods in which given a test point x_0 KNN finds K neighbors and then estimates the class probabilities as the fraction of neighbors which belong to a particular class.

$$Pr(Y = j|X = x_0) = \frac{1}{K} \sum_{i \in N_0} I(y_i = j) \tag{1.5}$$

Where N_0 are the nearest neighbors to the point x_o).

- Small K results in high variance
- Big K results in high bias
- K controls the trade off
- $\frac{1}{K}$ is a measurement of flexibility

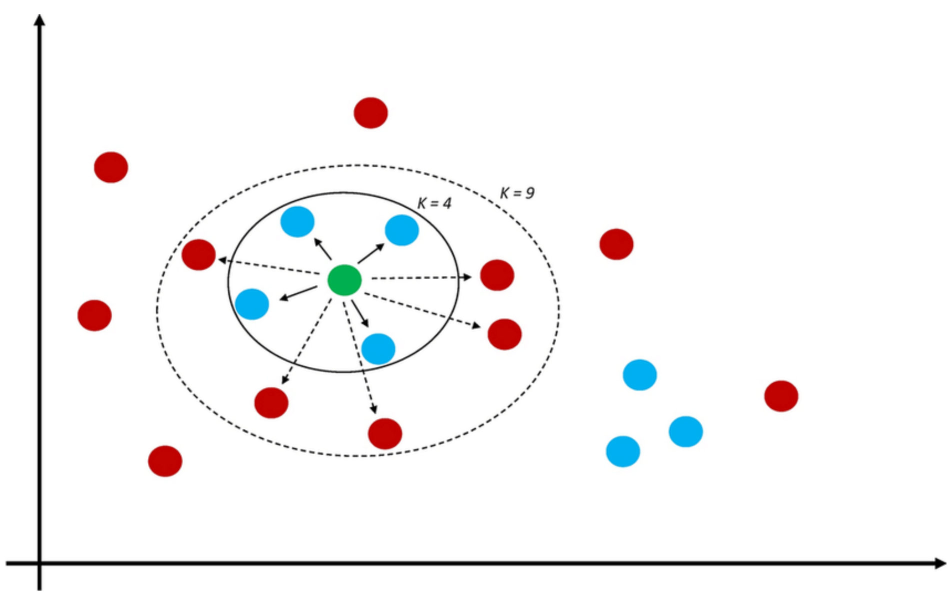


Figure 1.6: KNN

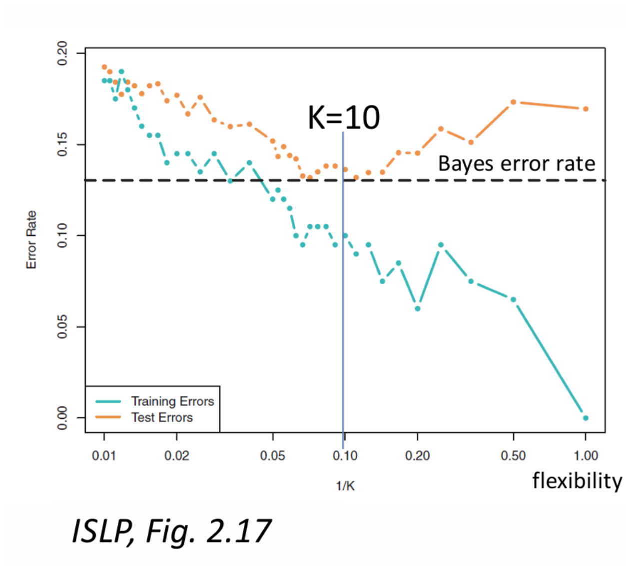


Figure 1.7: KNN Flexibility

5 Linear Regression

Definition 5.1: Linear Regression Model

Estimates the linear relationship between a scalar response (dependent variable) and one or more explanatory variables (regressor or independent variable). There is a closed form solution for the RSS optimization.

$$Y \approx \hat{Y} = \beta_0 + \sum_{i=0}^{n} \beta_{i+1} X_i \tag{1.6}$$

By minimizing RSS:

$$RSS = e_1^2 + \cdots + e_n^2 \text{ where } e_i = y_i - \hat{y}_i \tag{1.7}$$

5.1 Metrics

RSE is roughly speaking, the average amount that the response will deviate from the true regression line. It provides an absolute measure of lack of fit of the model to the data. In fact the RSE carries the same unit as Y. R^2 is the proportion of the variance explained by the model and hence it takes values between 0 and 1. **TSS** is the total sum of squares, which is the RSS if Y would always be predicted using the sample mean of Y (best possible predictor if we do not measure X). Assess accuracy of model with:

Definition 5.2: Model Accuracy

$$\begin{aligned} RSE &= \sqrt{\frac{1}{n-p-1} RSS} \\ TSS &= \sum (y_i - \bar{y})^2 \\ R^2 &= \frac{TSS - RSS}{TSS} = 1 - \frac{RSS}{TSS} \end{aligned} \tag{1.8}$$

In simple linear regression the following holds:

$$R^2 = Cor(X, Y)^2 = Cor(Y, X)^2 \tag{1.9}$$

In multivariate linear regression the following holds:

$$R^2 = Cor(Y, \hat{Y})^2 \tag{1.10}$$

We require a small variance of the estimate or equivalently a small standard error. Since σ^2 is unknown it has to be estimated from data. Note that σ is the standard deviation of each independent realization of Y.

$$\begin{aligned} Var(\hat{\mu}) = SE(\hat{\mu})^2 = \frac{\sigma^2}{n} \\ \text{where } \sigma^2 = Var(\epsilon) \end{aligned} \tag{1.11}$$

Variance σ^2 estimation using residual standard error (RSE):

Definition 5.3: Standard Error Estimation using RSE

$$\begin{aligned} \text{for } \sigma^2 = Var(\epsilon) \\ \implies SE \approx \hat{SE} = RSE = \sqrt{\frac{RSS}{n - n_\beta}} \end{aligned} \tag{1.12}$$

Under the Gaussian assumption confidence intervals can be computed using the estimated variance.

Definition 5.4: Confidence Intervals

$$Pr(a < \hat{b}_i < b) = 0.95 \implies \hat{\beta}_i \pm 2 \cdot \hat{SE}(\hat{\beta}_i) \tag{1.13}$$

5.2 T-Test

The T-test (with the T-statistic), is a tool for evaluating the means of one or two populations using hypothesis testing) Hypothesis tests using null hypothesis. We can test the relevance of coefficients using the T-test:

Definition 5.5: T-Test for Coefficients

$$\begin{aligned} H_0 : \text{There is no relation between } X \text{ and} \\ \text{or } b_i = 0 \\ H_a : \text{There is a relationship} \end{aligned} \tag{1.14}$$

$$\text{T-statistic} = \frac{\hat{\beta}_i}{\hat{SE}(\hat{\beta}_i)} \tag{1.15}$$

With $n - n_\beta$ degrees of freedom where p:

$$p = Pr(|T| > |t| \qquad |H_0) \tag{1.16}$$

P-Value is the probability that we realistically observe an absolute T-value equal or bigger to the one observed under the H_0.

5.3 F-Test

An F-test is any statistical test used to compare the variances of two samples or the ratio of variances between multiple samples. To check if there is any relationship between predictors and the dependent variable the F-test can be used:

Definition 5.6: F-Test

$$\begin{aligned} H_0 &: \beta_0 = \cdots = \beta_n = 0 \\ H_a &: \text{at least one } \beta_j \text{ is non-zero} \\ \text{using F-statistic} &: F = \frac{(TSS - RSS)/p}{RSS(n-p-1)} \\ \text{and} &: TSS = \sum (Y_i - \bar{y})^2,\ RSS = \sum (y_i - \hat{y})^2 \end{aligned} \tag{1.17}$$

F-statistic expresses the improvement of the model per parameter p in multiples of the residual variance. If the linear model assumptions are correct and H_a is true we expect the F-statistic on average to be greater than one.

One can also use a subset of q coefficients to test with a new hypothesis. We order the predictor variables so that the last q variables are the one to test for being zero. Then we fit a second model that uses all the predictor variables except the last q .

$$F = \frac{(RSS_0 - RSS)/q}{RSS/(n-p-1)} \tag{1.18}$$

Definition 5.7: Single Coefficient F-Test

When we leave out only one variable (q=1) then the F-statistic would tell the partial effect of adding that variable to the model. **It turns out, the F-statistic, if only one variable is left out, is identical to the squared T-statistic.**

5.4 Predictor Selection

- Forward selection
- Backward selection
  - Cannot be used if there are more predictor variables than training samples
- Mixed Selection

Part V

Parallel Computation

Part VI

Distributed Systems

Part VII

Databases

Part VIII

Security & Cryptography

Part IX

Software Engineering

CHAPTER 1

APPLICATION ARCHITECTURE

Definition 0.1: Software Architecture

The fundamental organization of a system is embodied in its components, their relationships to each other, and to the environment, and the principles guiding its design and evolution. (ISO/IEC/IEEE 42010, November 2011)

ARC42 Software Architecture Templates Fowler's Enterprise Architecture Fowler's GUI Architectures

1 Architecural Significant Requirements

1. The requirement is directly associated with high business value or business risk
2. The requirement is a concern of a particularly important stakeholder (for instance, the project sponsor or an external compliance auditor).
3. The requirement has runtime Quality-of-Service (QoS) characteristics (e.g., performance needs) that deviate from those already satisfied by the evolving architecture substantially.
4. The requirement causes new or deals with one or more existing external dependencies that have unpredictable, unreliable and/or uncontrollable behavior.
5. The requirement has a cross-cutting nature and therefore affects multiple parts of the system and their interactions; it may even have system-wide impact.
6. The requirement has a first-of-a-kind character: e.g., the team has never built a component before that satisfies this particular requirement.
7. The requirement has been troublesome and caused critical situations, budget overruns or client dissatisfaction in a previous project in a similar context.

1.1 NFR Catalogs and Taxonomies

Definition 1.1: SMART

Specific Targeting a particular area for improvement

Measurable Quantifying, or at least suggesting, an indicator of progress

Assignable Defining responsibility clearly

Realistic Outlining attainable results with available resources. Time-related: Including a timeline for expected results

The specific definitions of the words are not fixed.

The SMART criteria can be applied to NFR engineering.

- Specific: Which feature or part of the system should satisfy the requirement?
- Measurable: How can testers and other stakeholders find out whether the requirement is met (or not)? Is the requirement quantified?
- A, R, T are requirements engineering and project management concerns: Useful interpretation in our NFR context: Agreed Upon, Realistic, Time-Bound

Definition 1.2: FURPS+

- Functionality
- Usability
- Reliability
- Performance
- Supportability

Plus:

- Design constraints
- Implementation constraints
- Physical constrains
- Interface contraints

1.2 Quality Attribute Scenario (QAS)

Scenario for Order Management SOA at "T" (Business-to-Business), NFR #5		
Scenario Synopsis		Response Time in Web Channel: Virtual Service Provider (VSP) API
Business Goals		Drive down cost of operations by interacting with VSPs efficiently
Relevant Quality Attributes		Performance (response time), scalability
Scenario Components	Stimulus	Order placed by VSP (system or end user)
	Stimulus Source	External to system (see functional requirements and architecture overview diagram: "Create phone service", "Move phone service")
	Environment	At runtime, normal operational conditions: system up and running, not stressed (by unexpected workload, by an external attack, …)
	Artifact	All components and connectors on different logical layers and physical tiers required to process order (down to customer database in backend)
	Response	Orders are accepted immediately, order number (identifier) is returned
	Response Measure	Order acknowledgment is sent in 3 seconds (or less). This performance is desired but not guaranteed to VSPs in any Service Level Agreement (SLA).
Questions		Can errors be identified and handled properly in this timeframe?
Issues		Backend systems might not have sufficient processing power

Figure 1.1: QAS

Establish three measurable values (M) rather than a single measure that might not be realistic (R) and impossible to agree upon (A)

Response Time (per Business Activity, S)	Minimal Goal (Less Than)	Target (Within)	Outstanding (Within)
Order fully processed	2 weeks	24 hours	3 hours
Relocation processed	3 weeks	2 weeks	1 week
Technician appointment scheduled	2 days	1.5 days	1 day
Address validated	10 seconds	3 seconds	1 second
Billing system configured	1 week	3 days	1 day

Figure 1.2: Landing Zones

1.3 Twin Peaks

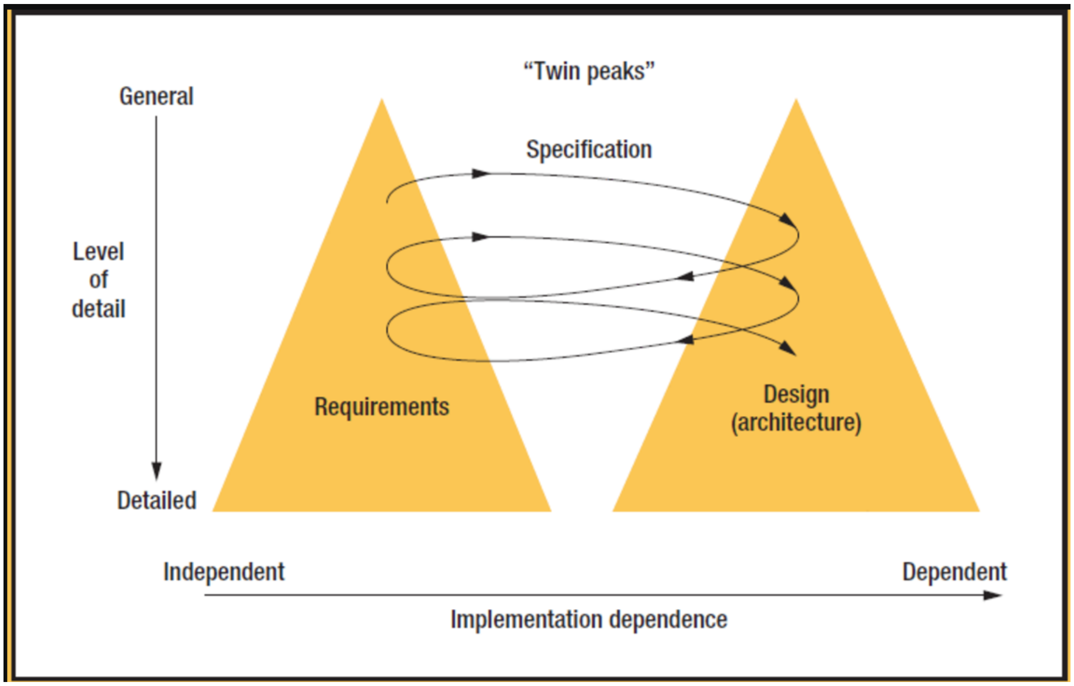


Figure 1.3: Twin Peak Design

2 Solution Strategy

Definition 2.1: Solution Strategy

fundamental decisions and solution strategies, that shape the system's architecture. These include:

- technology decisions
- decisions about the top-level decomposition of the system, e.g. usage of an architectural pattern or design pattern
- decisions on how to achieve key quality goals
- relevant organizational decisions, e.g. selecting a development process or delegating certain tasks to third parties.

ARC42 SS Template

2.1 Y-Template

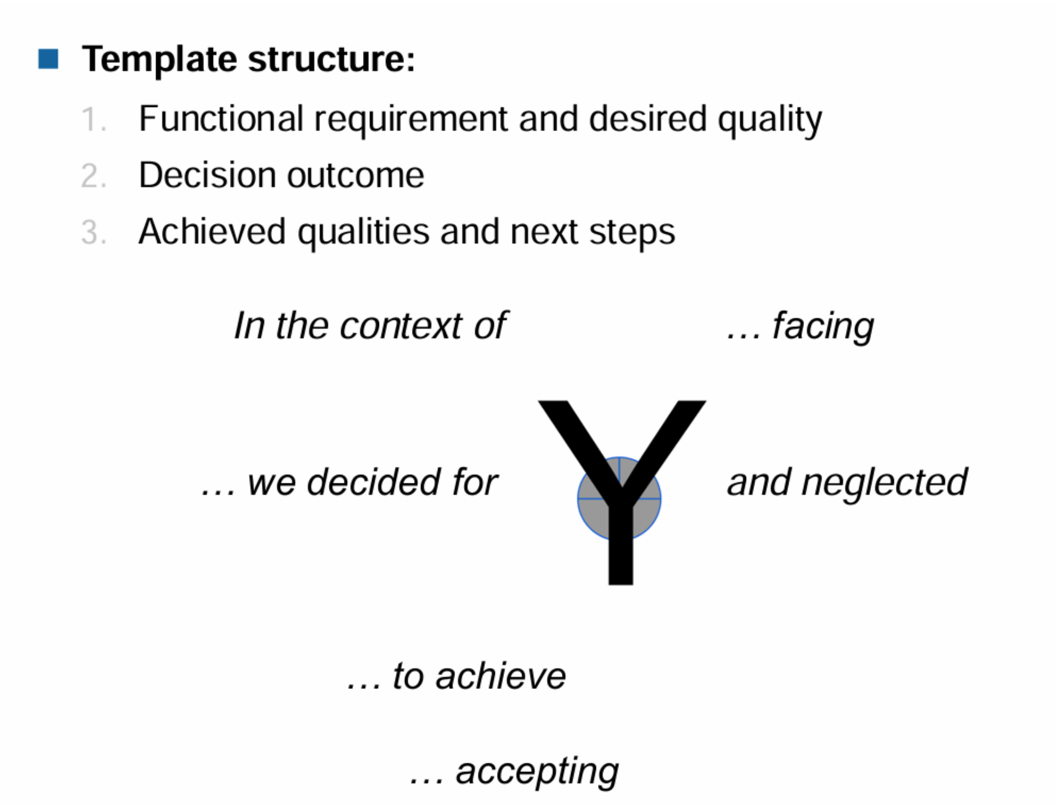


Figure 1.4: Y-Template for Architecture Design Decision

2.2 Architectural Decision Records

ADR Templates

2.3 Logical Layering

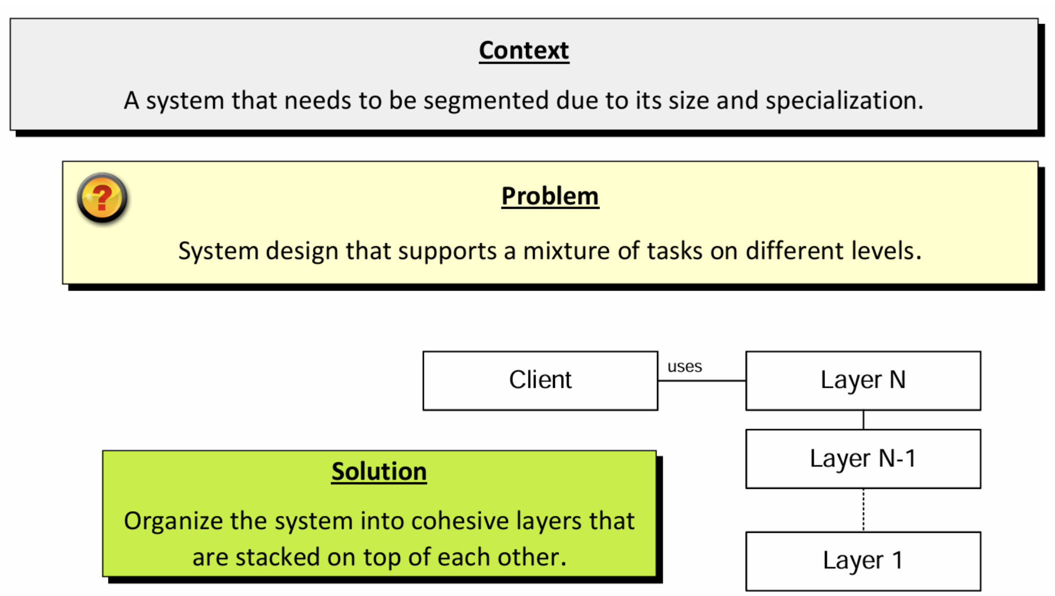


Figure 1.5: Logical Layering

- Layers separate concerns
- Tiers distribute workload

Presentation Layer: End users and external systems only talk to presentation layer. Rationale: isolation from backend. Presentation layer talks to business logic. Rationale: support multiple presentations of same logic

Business Logic: Business logic uses data access layer to communicate with database and backend systems. Which can be swapped in and out

The Layers pattern itself does not imply process/server boundary and any use of remoting is optional .

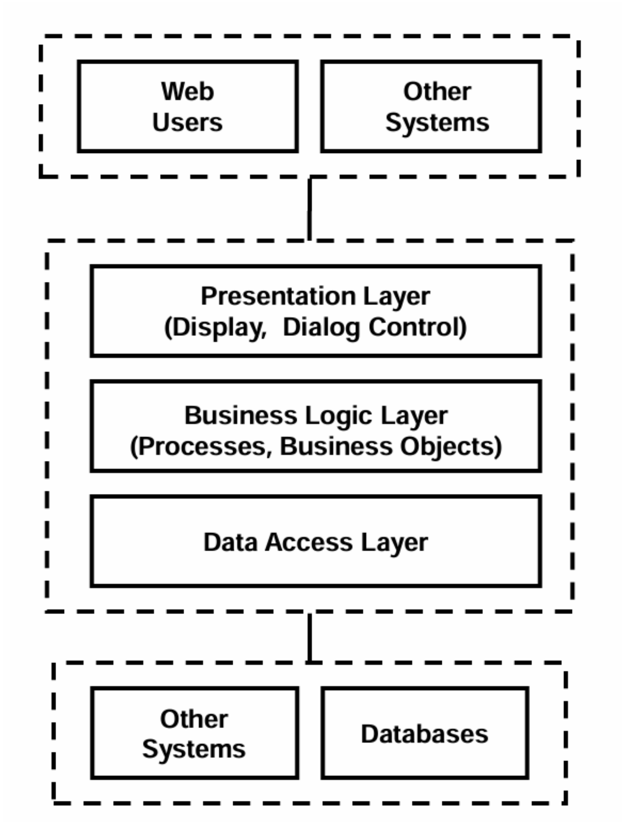


Figure 1.6: Logical Layers

Working with three primary logical layers

- Presentation
- Business logic
- Database (access)

With many options how to assign layers to tiers: Layer boundary or within layer, Single or multiple assignments.

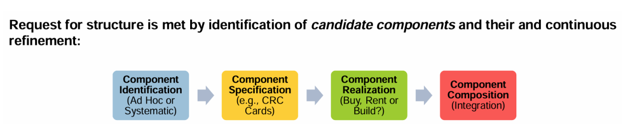


Figure 1.7: Candidate Components

2.4 Client Server Cuts (CSC)

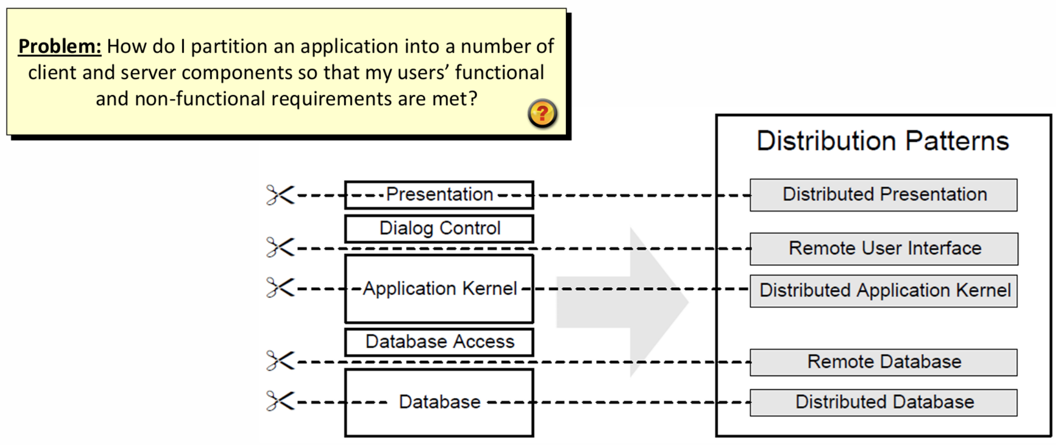


Figure 1.8: CSC

3 Big Architectural Decisions

1. Those with high arch. significance score
2. Those requiring financial investment; those with tough consequences
3. Those that take a long time to execute upon
4. Those with many or still unclear outgoing dependencies
5. Those that take a long time to make according to DoD (ecADR)
6. Those with a high level of abstraction
7. Those with problem/solution space outside of team's comfort zone

MADR Templates

4 Common Components

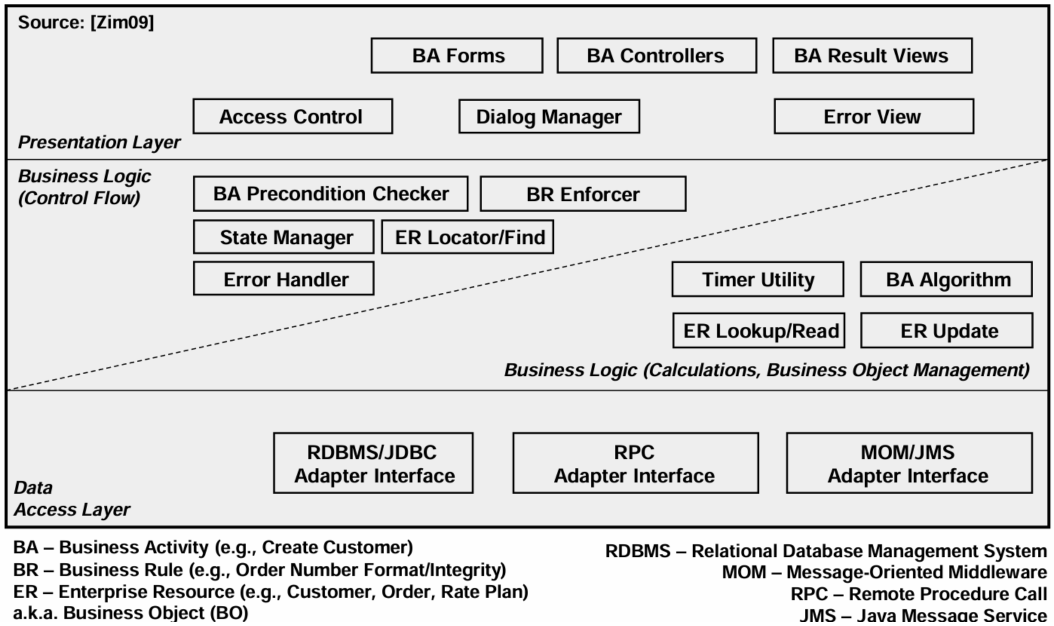


Figure 1.9: Common Components

5 Other Tools

6 C4 Model

Definition 6.1: C4 Model

The C4 model is a software architecture visualization model based on a set of hierarchical abstractions: software systems, containers, components, and code. It uses corresponding hierarchical diagrams—System Context, Container, Component, and Code—to describe these abstractions. The model is intentionally notation-independent and tooling-independent, focusing on clarity and adaptability rather than imposing specific methodologies or tools.

C4 model as a "light" alternative to UML.

1. First level of design (solution strategy): containers
2. Second level of design (refinement): components
3. Third level of design coded, not diagrammed (construction)

Context matters (the first C in C4), System Context Diagram (SCD) also present in other software engineering and architecture design methods

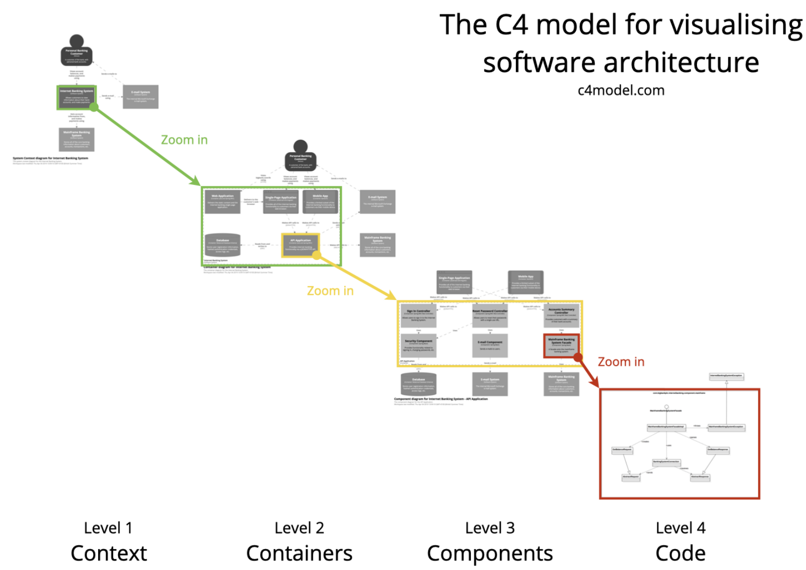


Figure 1.10: C4 Model Overview

Link to C4 Model Home Page

7 Example: Component Interaction Diagram

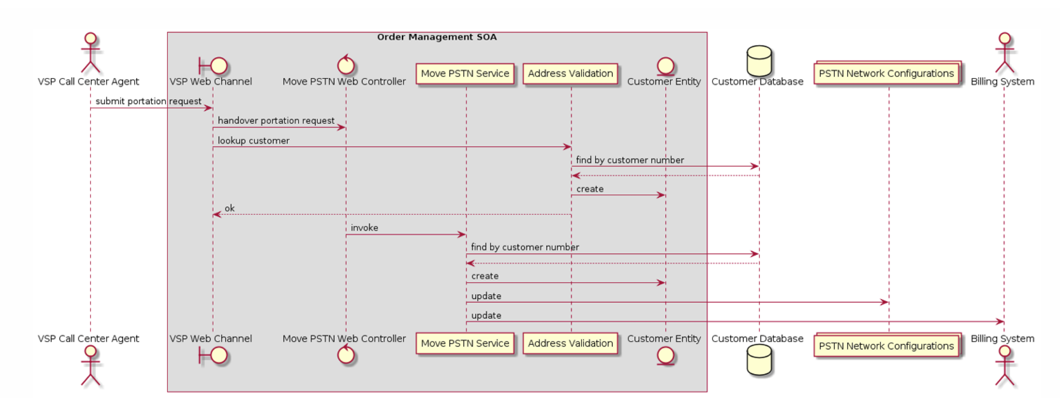


Figure 1.11: Example Component Interaction Diagram

8 Onion

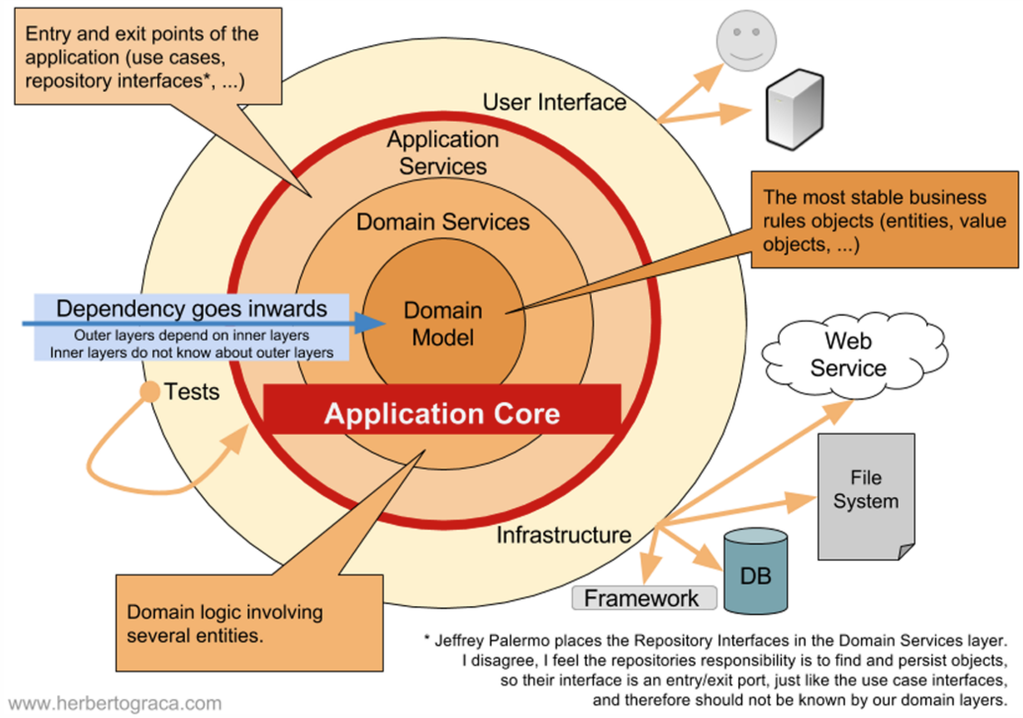


Figure 1.12: Onion Architecture

9 Story Splitting Flowchart

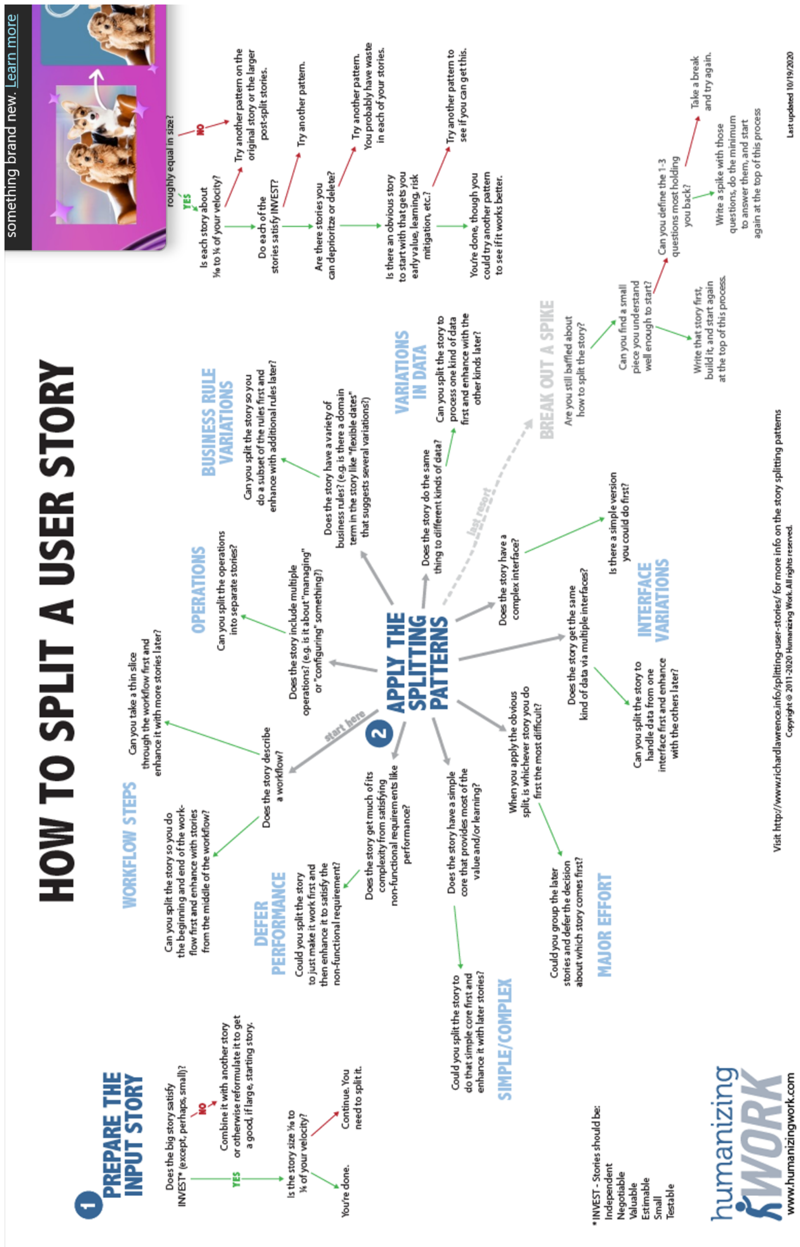


Figure 1.13: Story Splitting Flowchart

Splitting Pattern	Pres Layer Responsibility	Business Logic	Data Access Persistence Layer
Workflow Steps			
Operations			
Business Rule			
Data Variations			
Interface Variations			

Table 1.1: Story Splitting Example Table

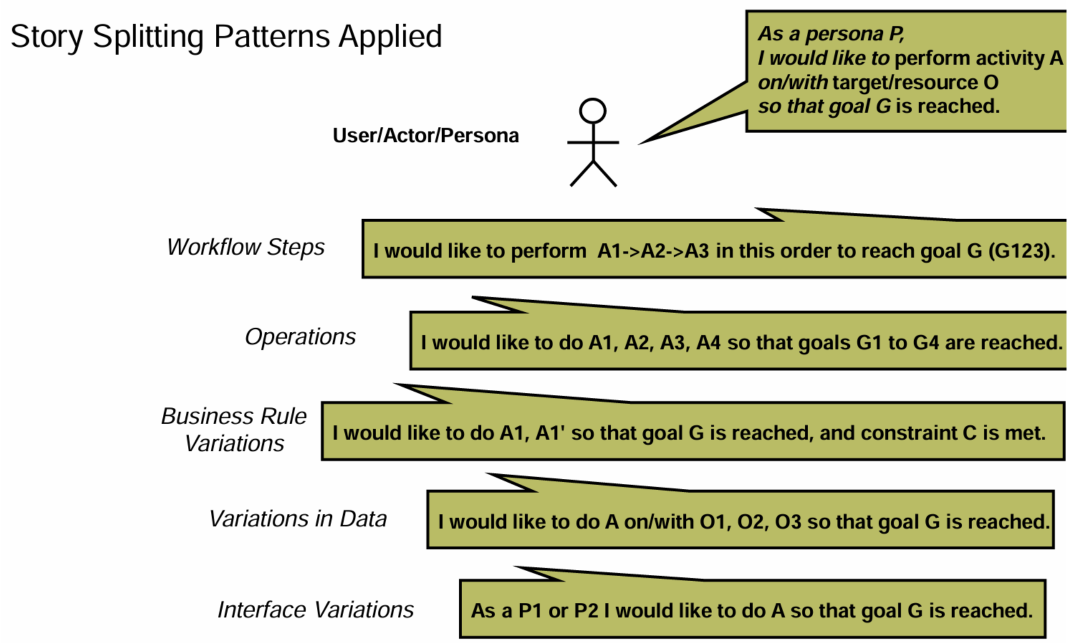


Figure 1.14: Story Splitting Example

10 AB CHAPTER??? 5-6-7??

Part X

Cheatsheets

BIBLIOGRAPHY

Black et al.: The Pricing of Options and Corporate Liabilities **BlackScholes**

Fischer Black and Myron Scholes. "The Pricing of Options and Corporate Liabilities". In: *Journal of Political Economy* 81.3 (1973), pp. 637–654.

Merton: Theory of Rational Option Pricing **Merton**

Robert Merton. "Theory of Rational Option Pricing". In: *The Bell Journal of Economics and Management Science* 4.1 (1973), pp. 141–183.

Hull: Options, Futures, and Other Derivatives **Hull**

John Hull. *Options, Futures, and Other Derivatives*. 9th ed. Pearson, 2017.