

**Schätzung der Erfolgsaussichten, der Dauer und
der Kosten von Softwareprojekten mit strengen
wahrscheinlichkeitstheoretischen Methoden**

Frank Padberg

1998

Fragen

Wie lange wird das Projekt dauern ?

Wieviel wird das Projekt kosten ?

Wie groß ist das Risiko, nicht rechtzeitig fertig zu werden ?

Wie groß ist das Risiko, das Budget zu überschreiten ?

Problem

der Verlauf eines Projekts

ist vorher nicht bekannt !

Idee

berechne die *Wahrscheinlichkeiten*

der einzelnen Projektverläufe

Literatur

Wolverton 74

Walston, Felix 77

Boehm 81

Albrecht 79

Putnam 76

Kitchenham, Taylor 84

Miyazaki, Mori 85

Kemerer 87

Kemerer 93

Wittig, Finnie 94

Srinivasan, Fisher 95

Shepperd, Schofield,

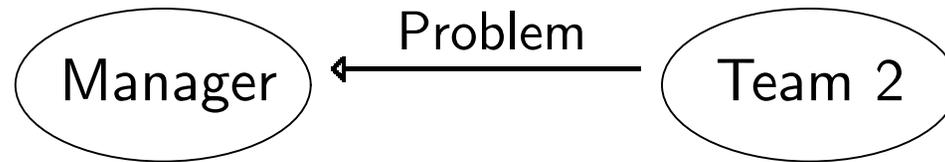
Kitchenham 96

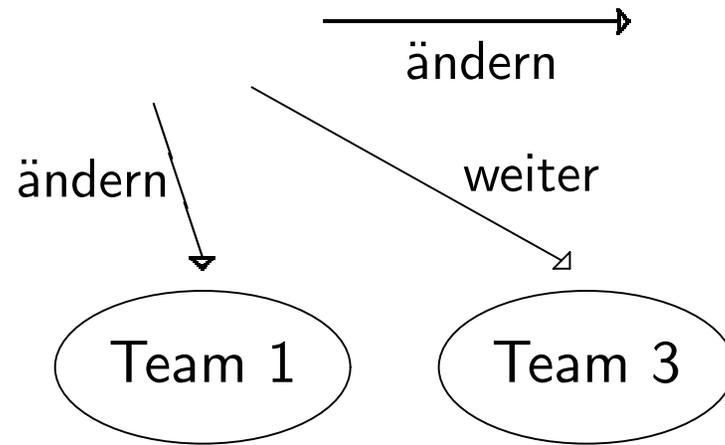
Gray, Mac Donell 97

Softwareprojekte

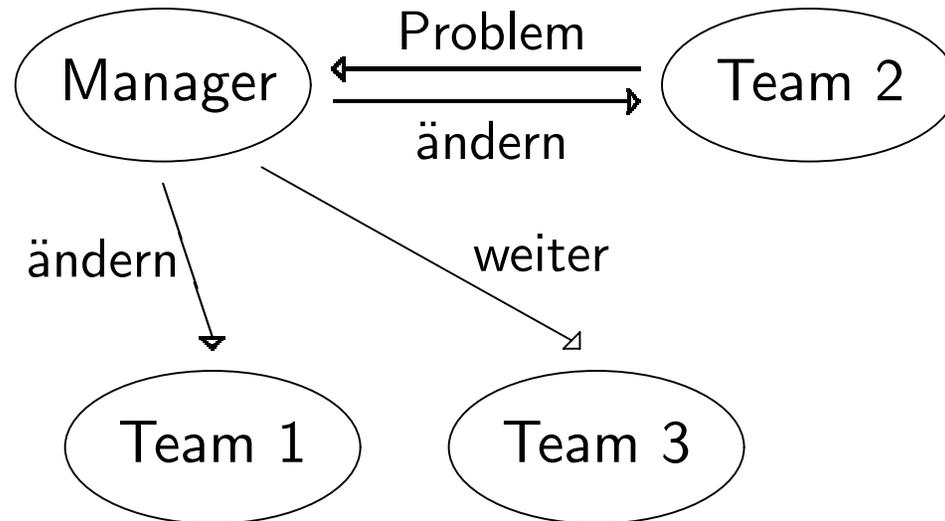
- mehrere Entwicklerteams und ein Projektmanager
- Entwurf ergibt Aufteilung in Teilaufgaben
- jede Teilaufgabe wird von einem Team bearbeitet
- die Teams arbeiten gleichzeitig
- Fortschritt eines Teams hängt von den anderen ab

Abhängigkeiten

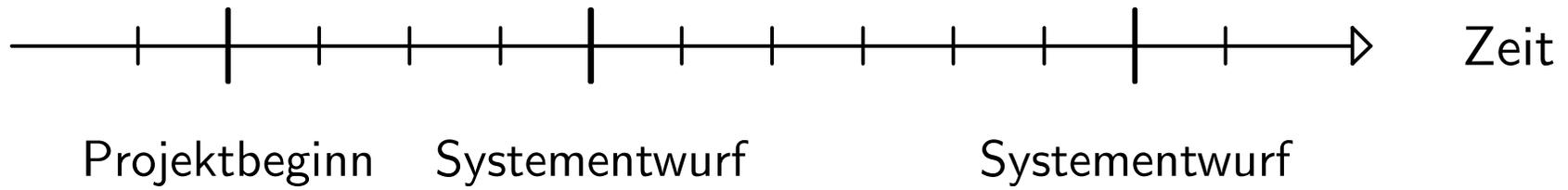




Abhängigkeiten



Projektverlauf



0

$\zeta(1)$

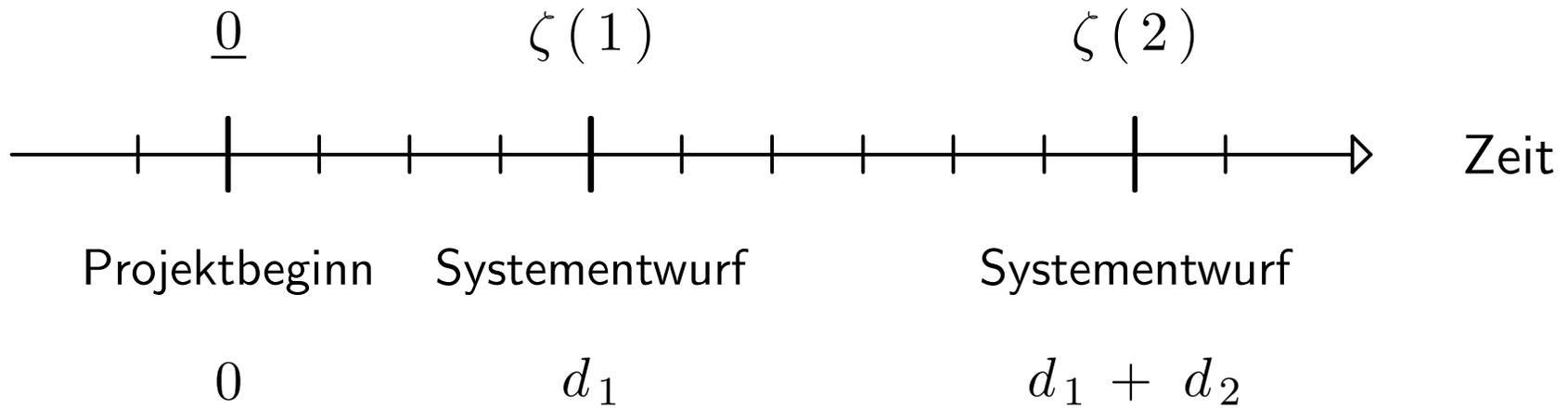
$\zeta(2)$

0

d_1

$d_1 + d_2$

Projektverlauf



Projektzustand

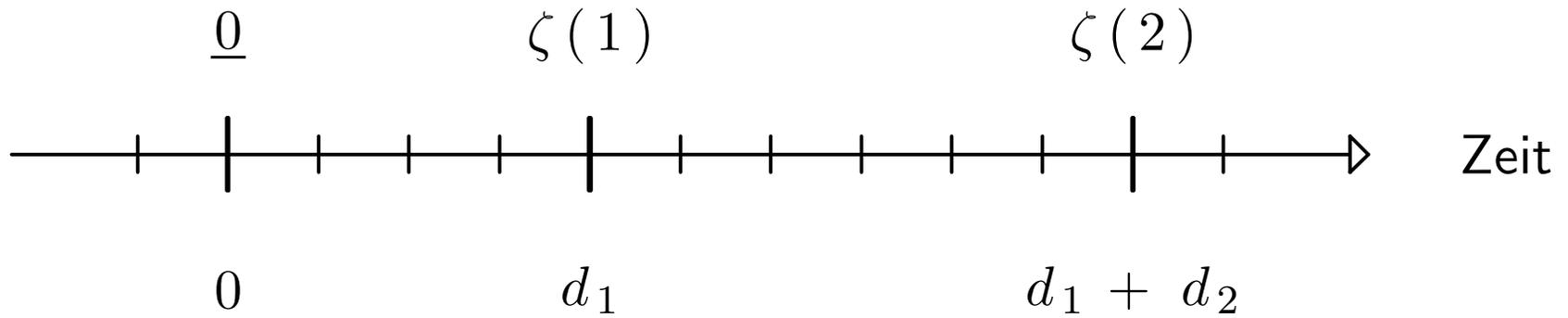
$$\zeta = \left(\zeta_1, \zeta_2, \dots, \zeta_N \right)$$

ζ_i – Zustand des i -ten Teams

Anzahl der Takte seit letzter Unterbrechung

Werte $0, 1, \dots, \infty$

Projektverlauf

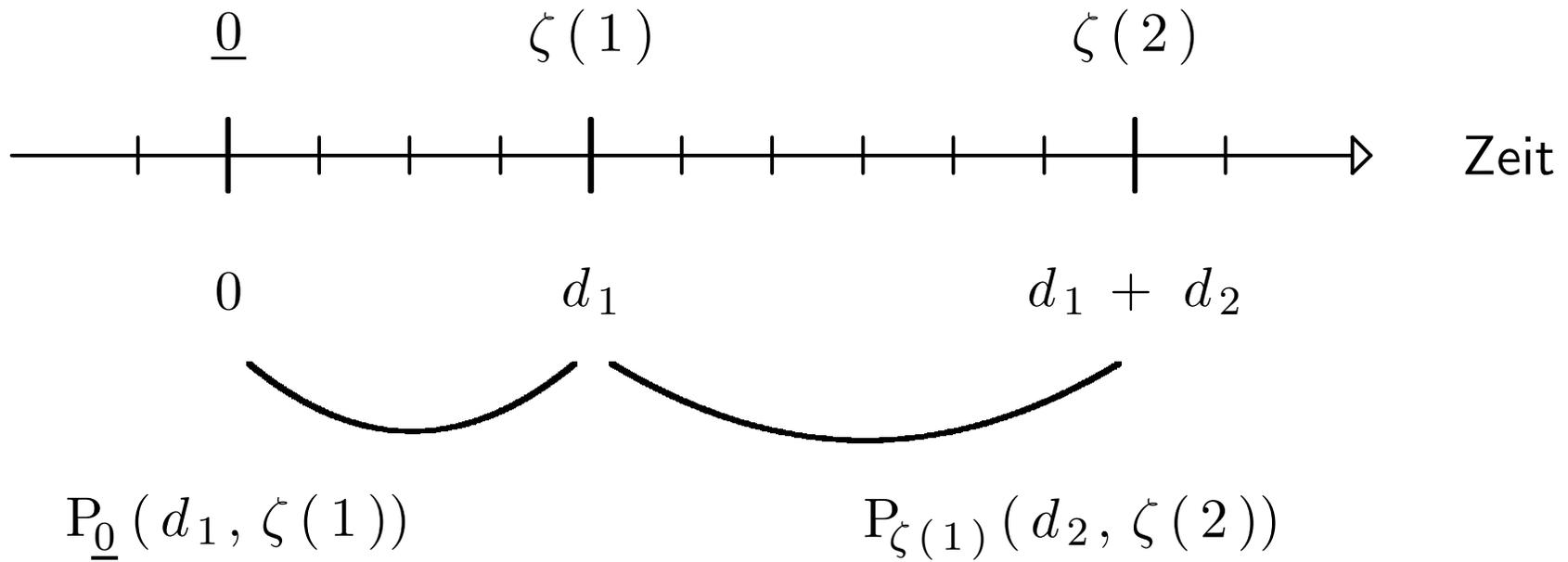


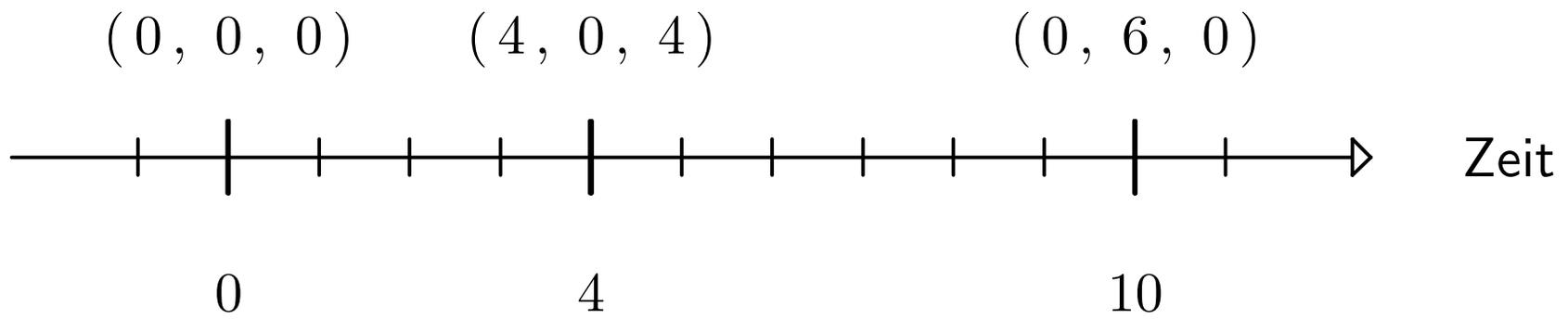


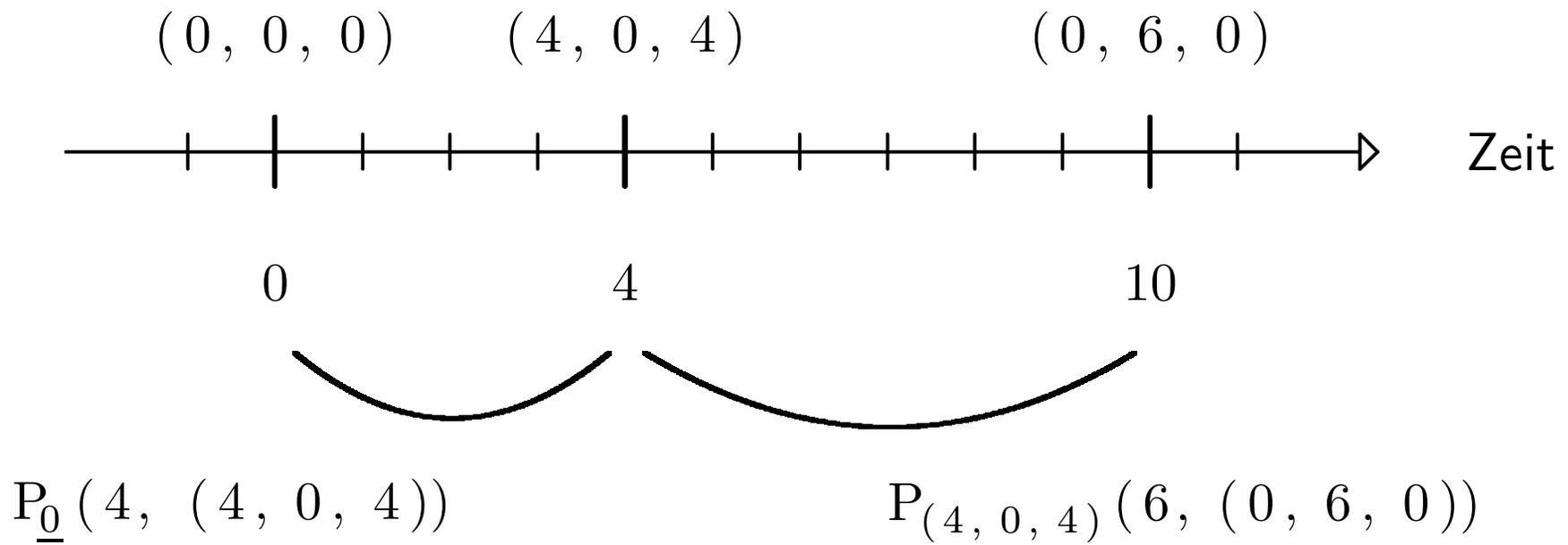
$P_{\underline{0}}(d_1, \zeta(1))$

$P_{\zeta(1)}(d_2, \zeta(2))$

Projektverlauf







Übergangswahrscheinlichkeiten

$$P_{\zeta}(d, \eta) = \sum_{v:d} P_{\zeta}(v) \cdot P_{\zeta}(\eta | v)$$

$P_{\zeta}(v)$ – Phase nimmt Verlauf v

$P_{\zeta}(\eta | v)$ – Verlauf v führt auf Folgezustand η

Daten

- statistische Daten über frühere Projekte
- Daten über den Entwurf der Software

Basiswahrscheinlichkeiten

$P(E_k^i)$ – Team i meldet nach genau k Takten ein Problem, sofern es nicht unterbrochen wird

$P(D_k^i)$ – Team i wird nach genau k Takten fertig, sofern es nicht unterbrochen wird

$\gamma(k)$ – genau k Takte nach einem Problem ist ein neuer Entwurf erstellt, sofern kein weiteres Problem auftritt

Statistische Rohdaten

bei laufenden Projekten regelmäßig aufzeichnen ...

- ... welche Teams fertig geworden sind
- ... welche Teams ein Problem gemeldet haben
- ... ob es einen neuen Entwurf gegeben hat
- ... wenn ja, welche Teams betroffen gewesen sind

Einflußgrößen

- Komplexität der Teilaufgabe
- verwendete Entwicklungstechniken
- Erfahrung der Entwickler

Faktoren in $P_{\zeta}(v)$

Team wird fertig: $P(D_{\zeta_i}^i + l_{\mu} \mid B_{\zeta_i}^i)$

Team meldet ein Problem: $P(E_{\zeta_i}^i + k_{\nu} \mid B_{\zeta_i}^i)$

Team meldet sich nicht: $P(B_{\zeta_i}^i + d \mid B_{\zeta_i}^i)$

kein Entwurf zwischendurch: $1 - \Gamma(k_{\nu+1} - k_{\nu} - 1)$

Entwurf am Phasenende: $\gamma(d - k_n)$

Verlauf paßt zum Anfangszustand: $1 / 0$

Abhängigkeitsgrade

$\alpha(K, X)$ – genau die Teams aus X sind von Änderungen betroffen, wenn genau die Teams aus K Probleme melden

Folgezustand

Team ist fertig: $\eta_i = \infty$

Team hat durchgearbeitet: $\eta_i = \zeta_i + d$

Team ist betroffen: $\eta_i = 0$

Faktoren in $\mathbb{P}_\zeta(\eta \mid v)$

betroffene Teams: $\alpha(K, \{i \mid \eta_i = 0\})$

Folgezustand paßt zu Verlauf und Anfangszustand: $1 / 0$

Vereinfachende Annahmen

- Anzahl der Teams ändert sich nicht
- Teams beginnen gleichzeitig
- Auftragstellung ändert sich nicht
- Ausmaß der Abhängigkeiten ist stabil
- Schwierigkeitsgrad der Teilaufgaben ist stabil

Beweis

- für jeden Zustand ζ gilt:

$$\sum_{d, \eta} P_{\zeta}(d, \eta) = 1$$

- für jeden Zustand ζ gilt:

$$\sum_v P_{\zeta}(v) = 1$$

Zeitschätzung

$$\text{Dauer: } f(\omega) = \sum_{j=1}^{|\omega|} d_j$$

$$\text{Erfolgskurve: } \varphi(x) = \sum_{\{\omega: f(\omega) = x\}} P(\omega)$$

$$\text{Erfolgsaussichten: } \Phi(y) = \sum_{x \leq y} \varphi(x)$$

$$\text{Schätzwert: } E_{dauer} = \sum_{x \leq x_0} x \cdot \varphi(x) + x_0 \cdot (1 - \Phi(x_0))$$

Kostenschätzung

$$\text{Kosten: } g(\omega) = \sum_{j=1}^{|\omega|} c_j$$

$$\text{Kostenkurve: } \psi(z) = \sum_{\{\omega: f(\omega) \leq x_0, g(\omega) = z\}} P(\omega)$$

$$\text{Erfolgsaussichten: } \Psi(y) = \sum_{z \leq y} \psi(z)$$

$$\text{Schätzwert: } E_{\text{kosten}} = \sum_{z \leq z_0} z \cdot \psi(z) + z_0 \cdot (1 - \Psi(z_0))$$

Personalbedarf

$3 \cdot x$	$g = 3 \cdot x$		$3 \cdot x$	$g = 3 \cdot x$	
	$f \leq x_0$	$f = x$		$f \leq x_0$	$f = x$
3	0.0	0.0	21	4.6	3.1
6	0.5	0.5	24	4.3	2.7
9	4.8	4.7	27	3.4	2.1
12	9.6	8.6	30	2.9	1.6
15	9.1	7.6	33	2.3	1.2
18	6.3	4.3	36	2.0	0.9

Risikoabschätzung

Erfolgsaussichten: $\Psi(y) = \sum_{z \leq y} \psi(z)$

Schwankungsbreiten:

$$P(|f - E_{dauer}| \leq y) \approx$$

$$\Phi(E_{dauer} + y) - \Phi(E_{dauer} - y)$$

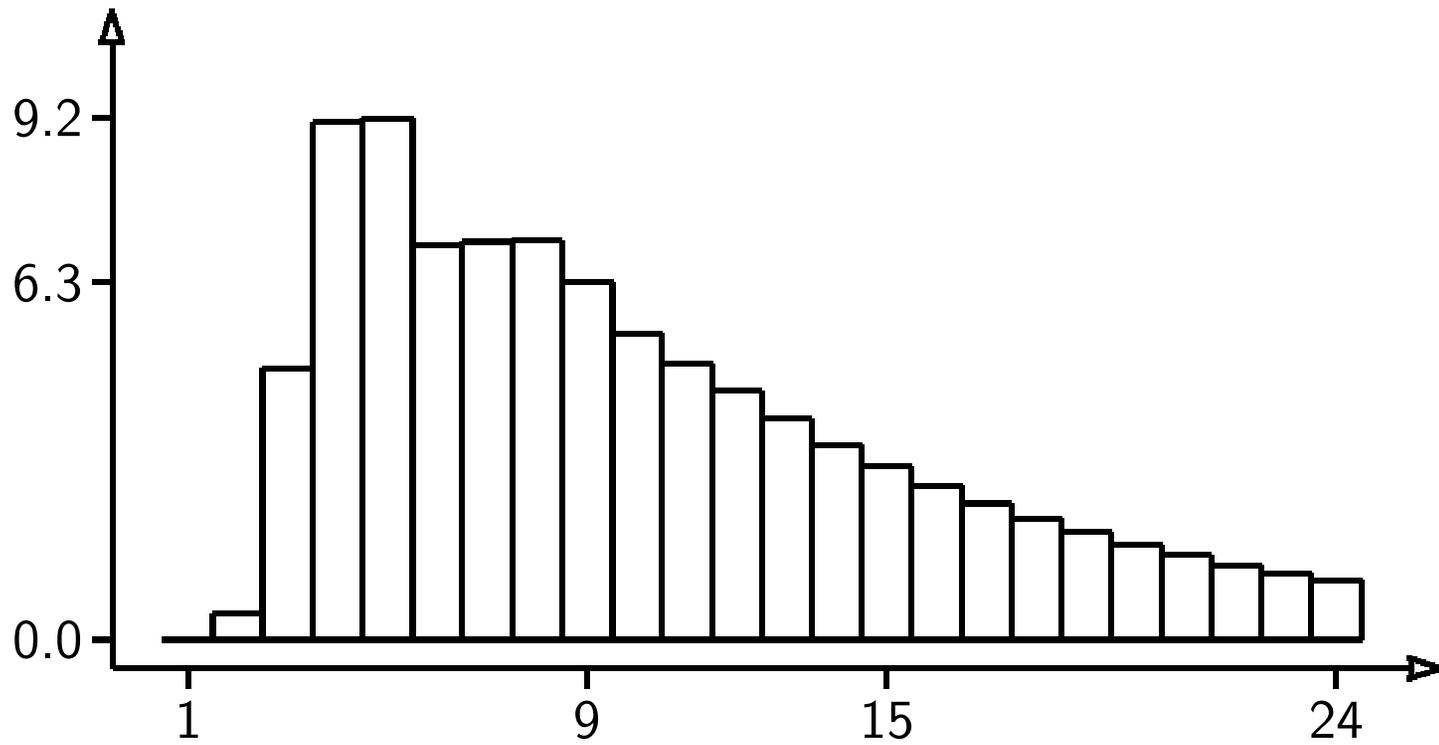
$$P(|g - E_{kosten}| \leq y, f \leq x_0) \approx$$

$$\Psi(E_{kosten} + y) - \Psi(E_{kosten} - y)$$

Erfolgskurve

Erfolgschance

- Prozent -



Zeit
- Monate -

Schätzwert : 11.1

Chancen : 92 %

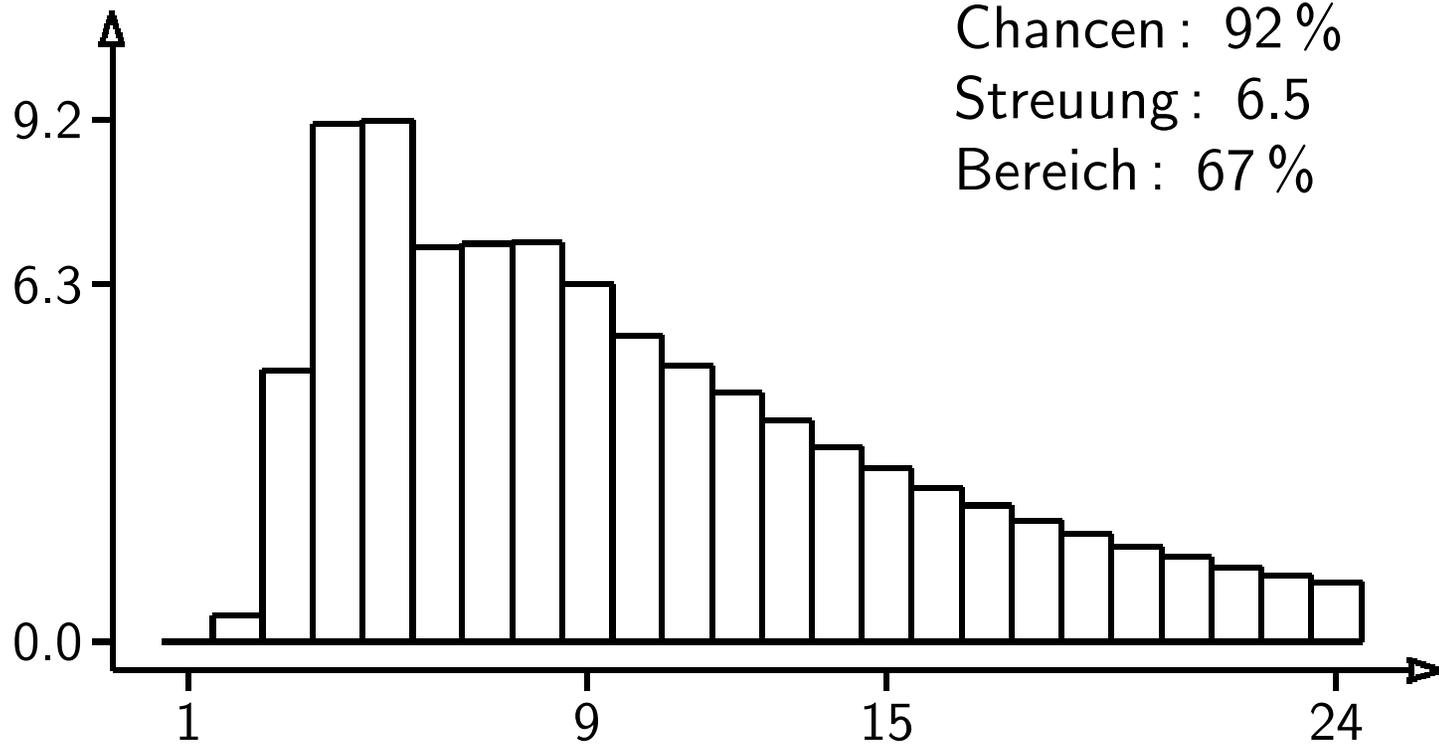
Streuung : 6.5

Bereich : 67 %

Erfolgskurve

Erfolgschance

- Prozent -

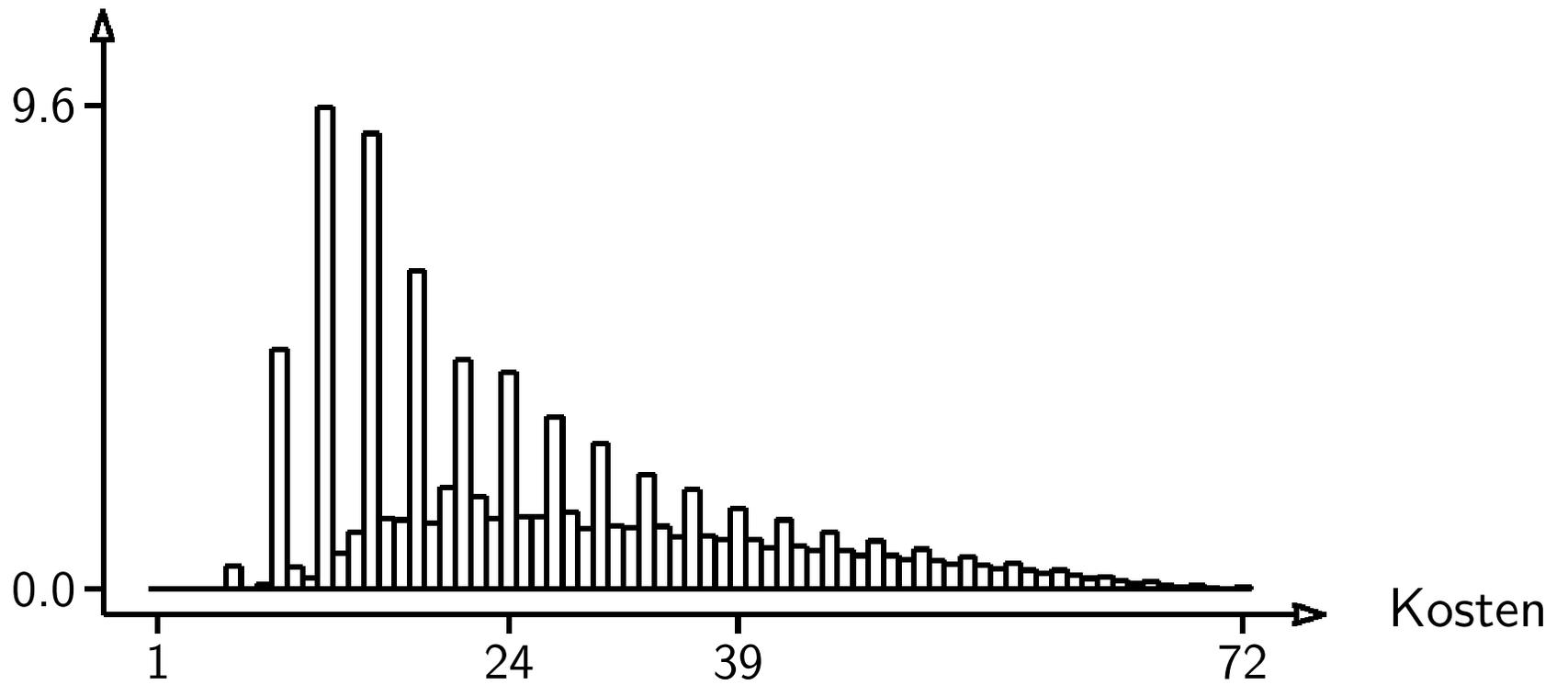


Zeit
- Monate -

Kostenkurve

Wahrscheinlichkeit

- Prozent -



Schätzwert : 29.9

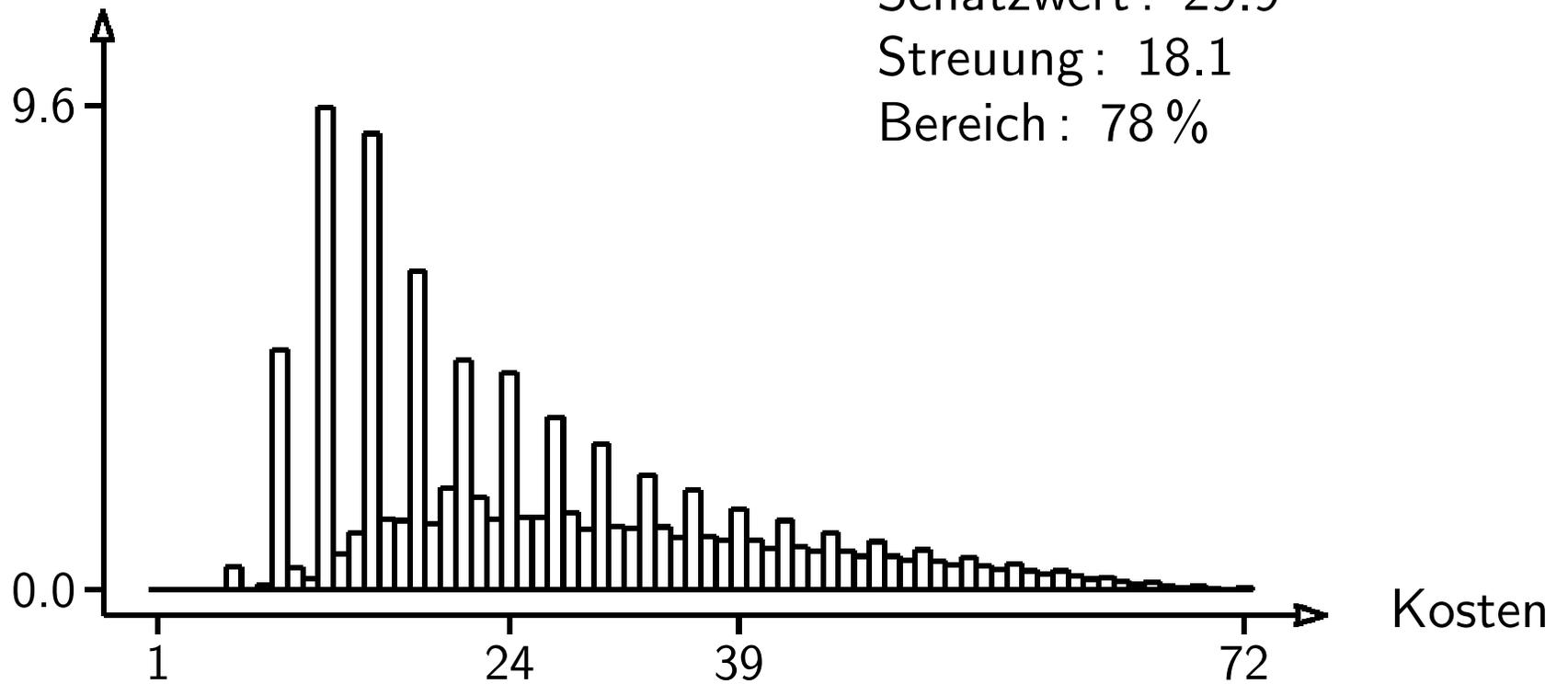
Streuung : 18.1

Bereich : 78 %

Kostenkurve

Wahrscheinlichkeit

- Prozent -



k	D_k^1	E_k^1	D_k^2	E_k^2	D_k^3	E_k^3
1	0.0	10.0	10.0	30.0	10.0	5.0
2	9.0	9.0	18.0	12.0	8.5	8.5
3	21.6	7.2	15.0	6.0	20.4	20.4
4	21.6	4.3	3.6	3.6	15.0	2.7
5	12.1	1.7	0.5	1.1	7.1	1.0
6	3.1	0.4	0.1	0.1	1.3	0.1

$$\gamma(1) = 100.0$$

$\alpha (K, X)$	1	2	3	1, 2	1, 3	2, 3	1, 2, 3
1	25	0	0	25	25	0	25
2	0	25	0	25	0	25	25
3	0	0	25	0	25	25	25
1, 2	0	0	0	50	0	0	50
1, 3	0	0	0	0	50	0	50
2, 3	0	0	0	0	0	50	50
1, 2, 3	0	0	0	0	0	0	100

Antworten

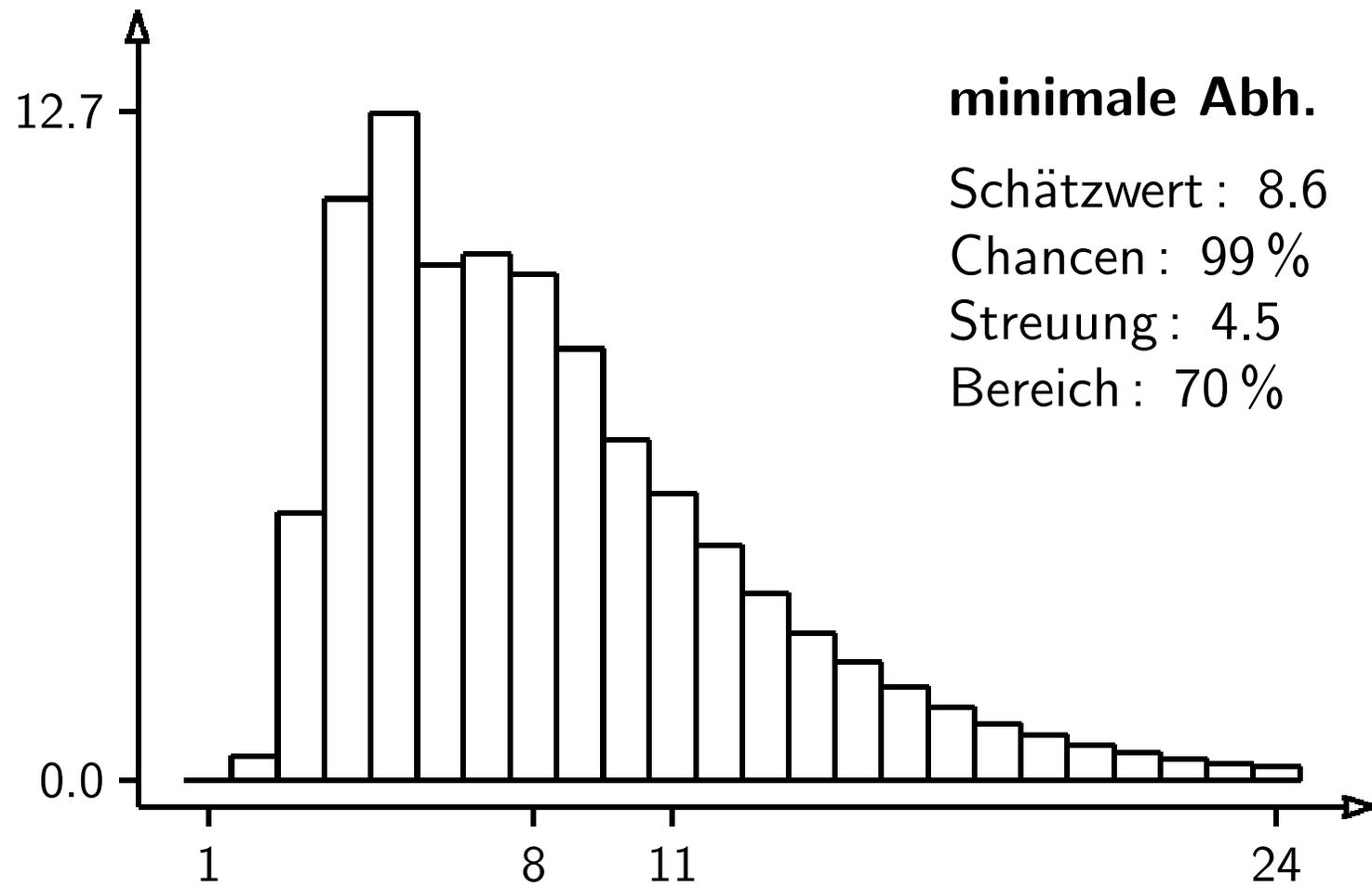
- ✓ Zeitschätzungen
- ✓ Kostenschätzungen
- ✓ Risikoabschätzungen

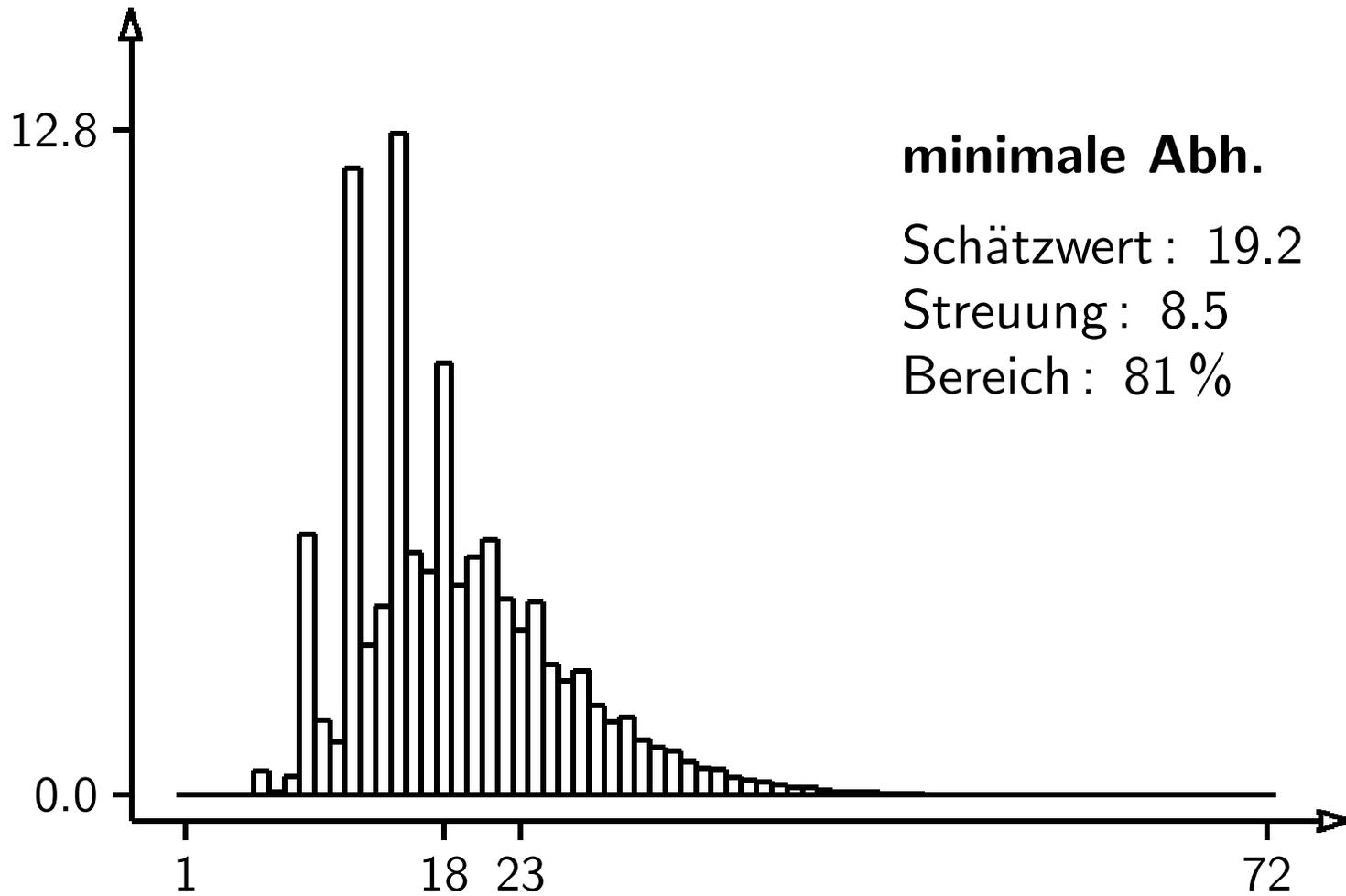
Ausblick

- reale Daten
- genauere Modellierung
- weitere Anwendungen

Modellverhalten

		α min	α max	W \uparrow	W \downarrow
Erfolgschance	92 %	99 %	79 %	96 %	89 %
Zeitschätzung	11.1	8.6	13.5	9.9	11.8
- Streuung	6.5	4.5	7.6	5.7	6.8
- Bereich	67 %	70 %	54 %	68 %	65 %
Kostenschätzung	29.9	19.2	40.4	26.8	32.0
- Streuung	18.1	8.5	22.7	15.5	19.6
- Bereich	78 %	81 %	54 %	82 %	67 %





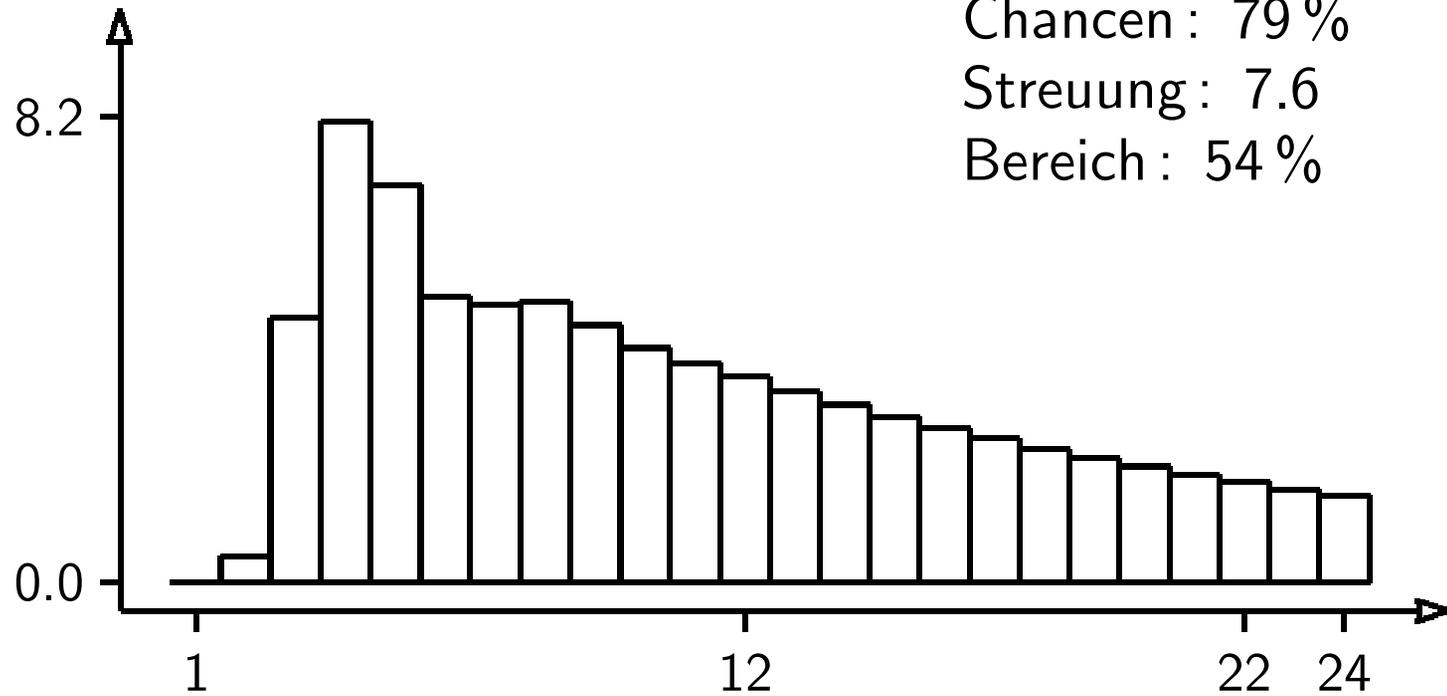
maximale Abh.

Schätzwert : 13.5

Chancen : 79 %

Streuung : 7.6

Bereich : 54 %

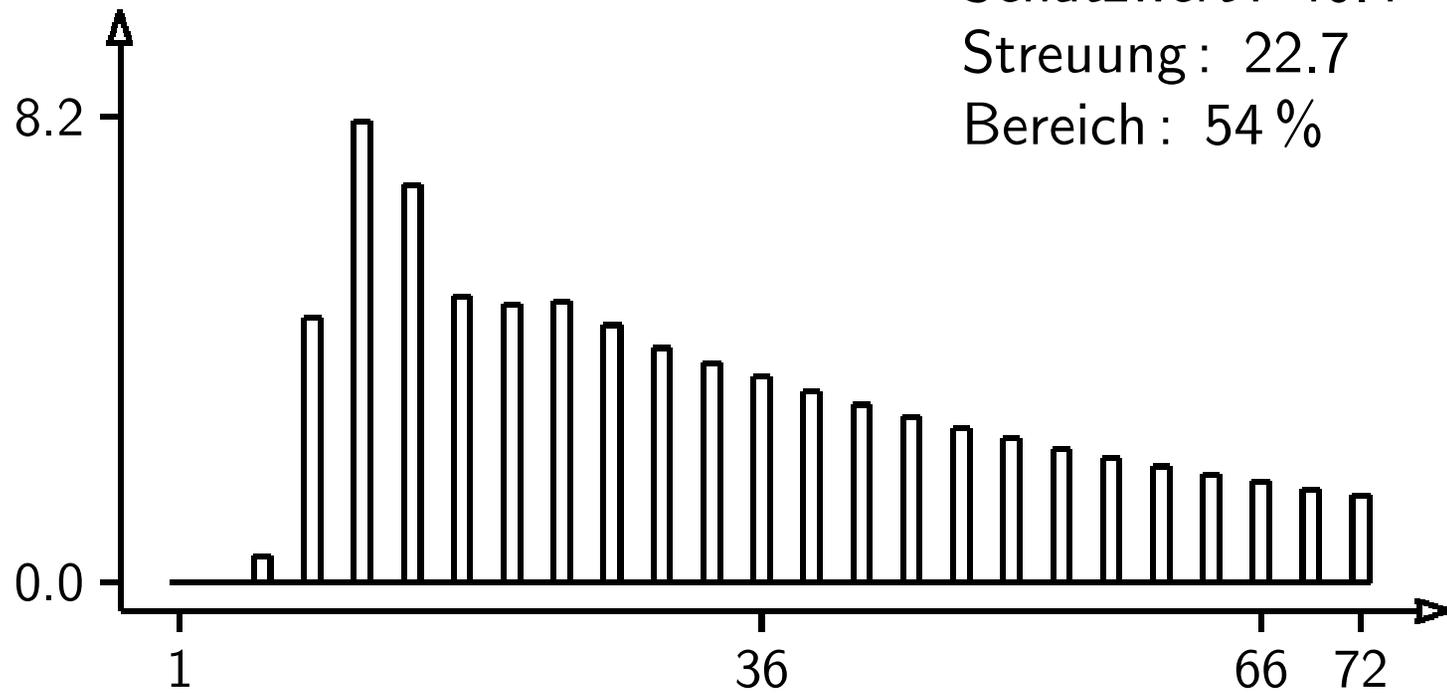


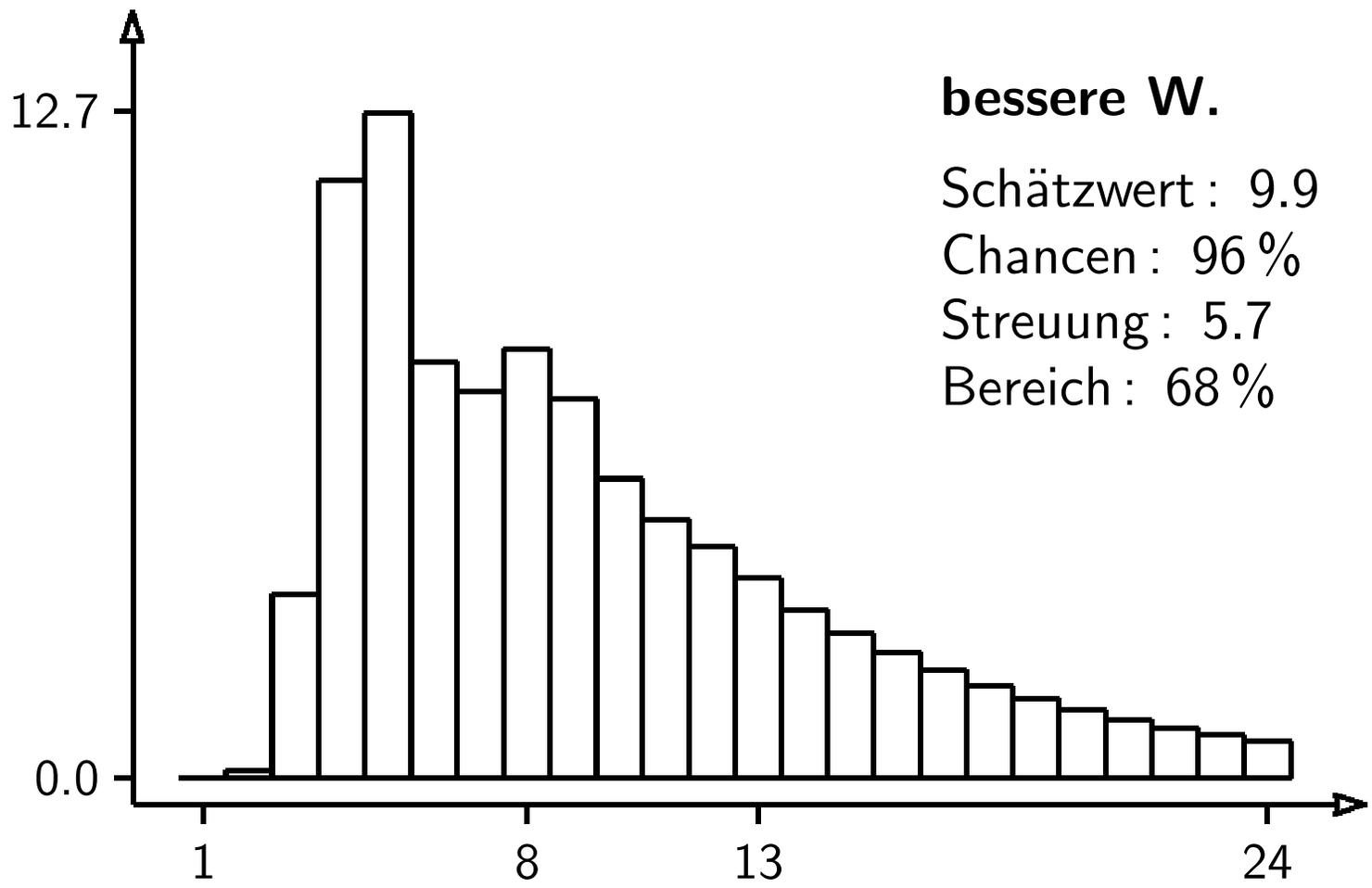
maximale Abh.

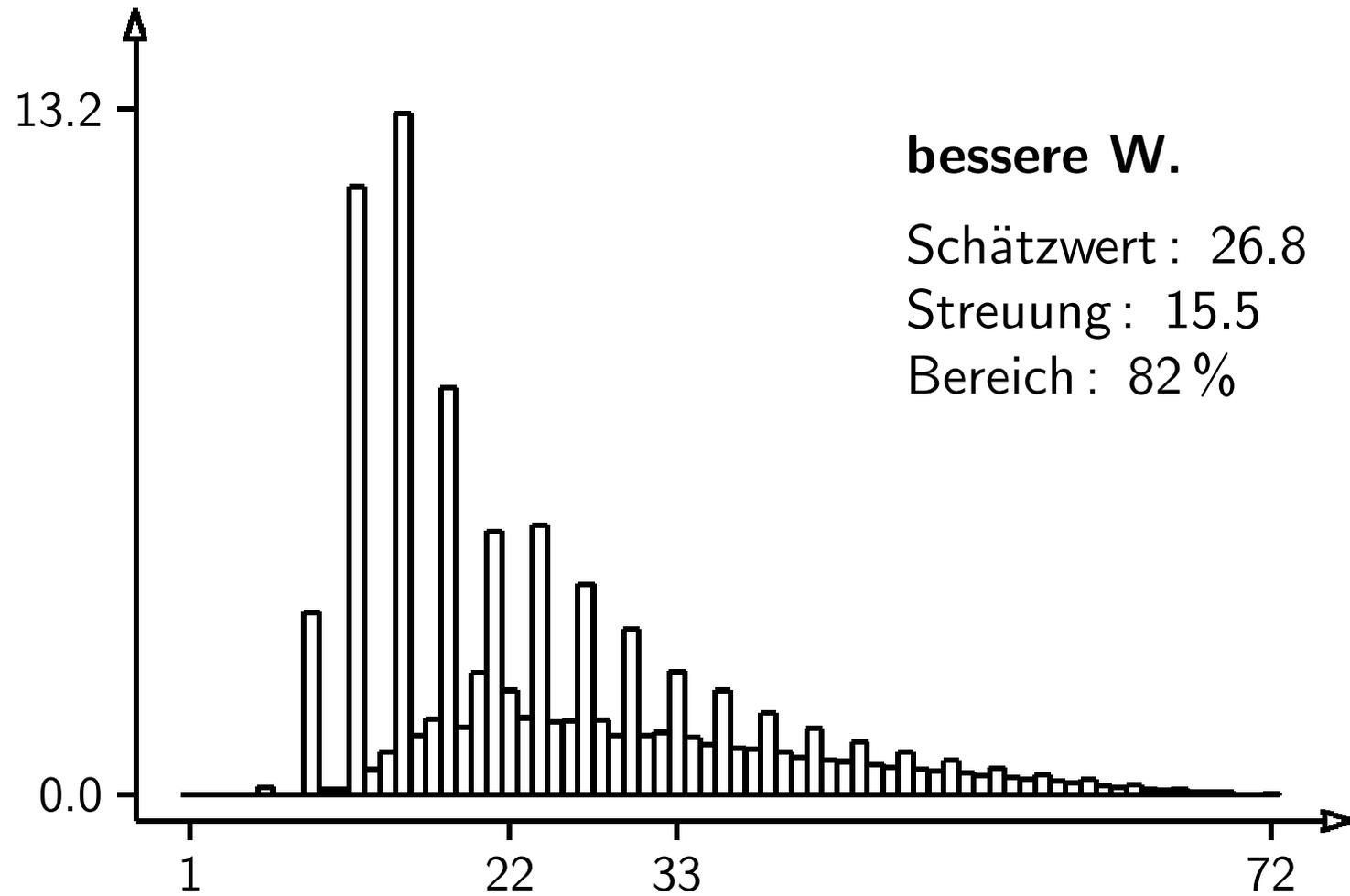
Schätzwert : 40.4

Streuung : 22.7

Bereich : 54 %







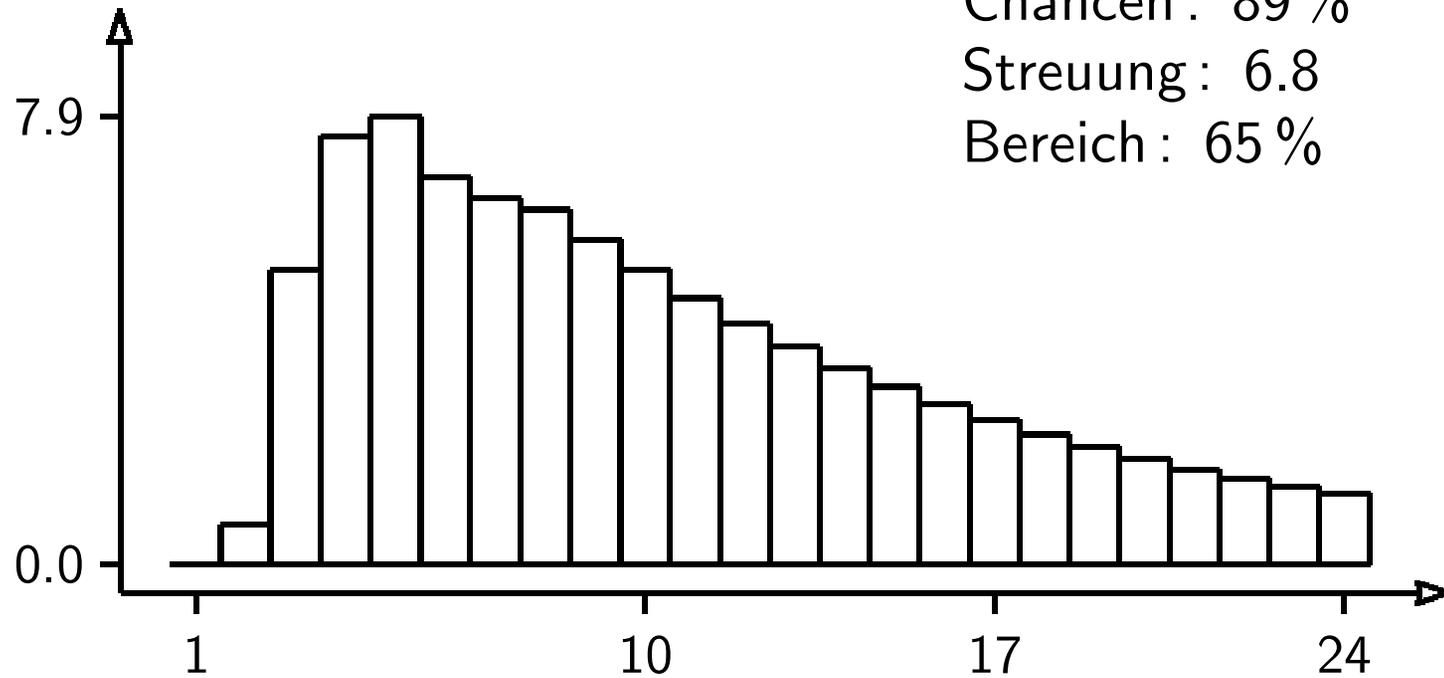
schlechtere W.

Schätzwert : 11.8

Chancen : 89 %

Streuung : 6.8

Bereich : 65 %

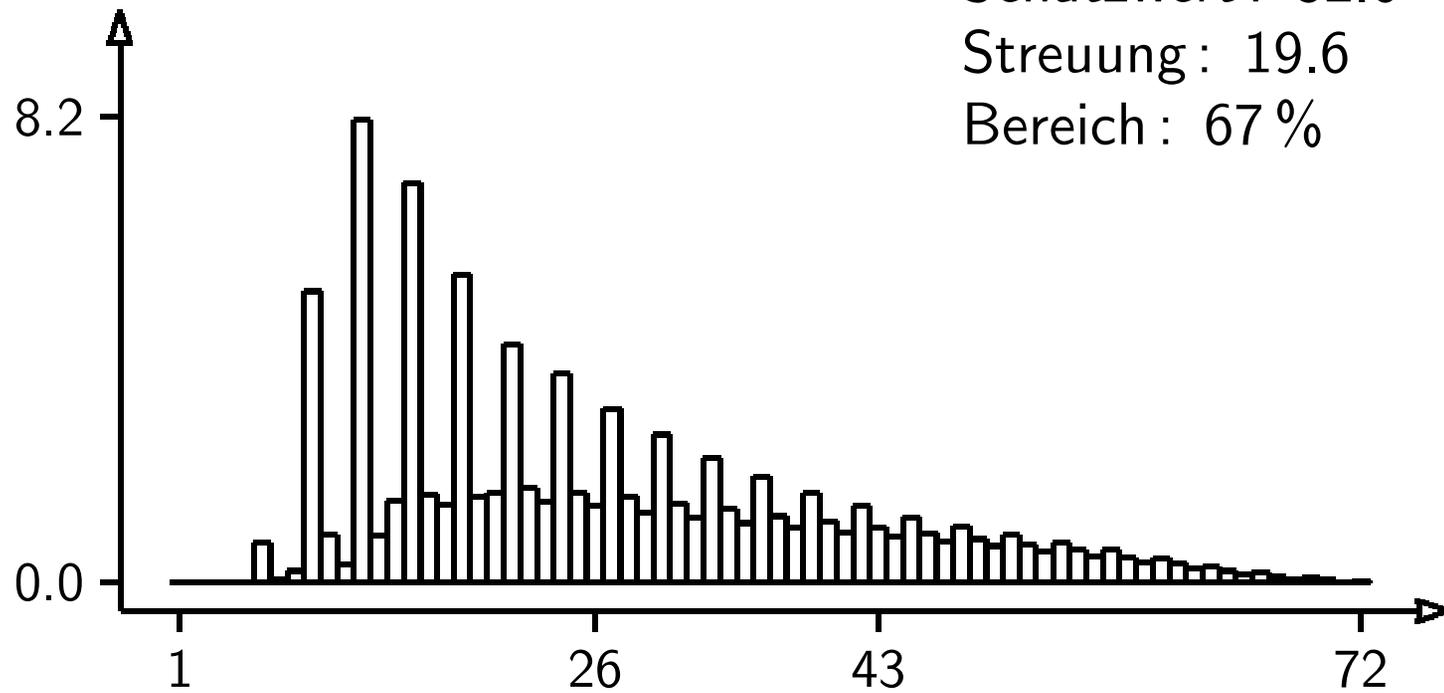


schlechtere W.

Schätzwert : 32.0

Streuung : 19.6

Bereich : 67 %



Rechenaufwand

Teams	3	4	5	6
Zeilen	74	396	867	8772
Einträge	3768	48 120	259 086	6 225 408
Rechenzeit – Matrix	0 : 01	0 : 06	2 : 30	119 : 10
Rechenzeit – Kurven	0 : 02	0 : 21	2 : 27	117 : 28

Simulation

- würfle von Takt zu Takt den Fortschritt der Teams aus:
 $P(E_{\zeta_i+k+1}^i \mid B_{\zeta_i+k}^i)$ und $P(D_{\zeta_i+k+1}^i \mid B_{\zeta_i+k}^i)$
- würfle nach Problemen von Takt zu Takt die Erstellung des neuen Entwurfs aus: $\gamma(j+1 \mid j)$
- würfle am Ende der Phase die betroffenen Teams aus:
 $\alpha(K, X)$

Simulationsergebnisse

Teams	Chance		Dauer		Streuung	
	ber.	sim.	ber.	sim.	ber.	sim.
3	91.8	91.7	11.1	11.1	6.5	6.5
4	85.0	84.7	13.2	13.2	6.8	6.9
5	74.5	74.7	15.3	15.3	7.0	7.0
6	65.8	65.6	16.9	16.9	6.9	6.8

Simulationsaufwand

Teams	Runden	MRE	Rechenzeit
3	30.000	2.4 %	0 : 20
4	40.000	2.8 %	0 : 38
5	50.000	2.5 %	1 : 07
6	60.000	2.1 %	1 : 44

N = 3 R = 30.000

	ber.	sim.		ber.	sim.		ber.	sim.	
1	0.00	0.00		9	6.29	6.36	17	2.41	2.46
2	0.46	0.46		10	5.40	5.18	18	2.13	2.05
3	4.78	4.85		11	4.86	4.78	19	1.89	1.88
4	9.12	9.13		12	4.38	4.55	20	1.68	1.65
5	9.17	9.35		13	3.89	4.00	21	1.49	1.56
6	6.97	7.04		14	3.43	3.50	22	1.32	1.36
7	7.02	7.00		15	3.05	2.88	23	1.17	1.09
8	7.03	6.84		16	2.71	2.66	24	1.04	1.01

N = 4 R = 40.000

	ber.	sim.		ber.	sim.		ber.	sim.	
1	0.00	0.00		9	6.06	6.26	17	2.88	2.87
2	0.04	0.05		10	5.37	5.32	18	2.62	2.57
3	1.54	1.63		11	4.92	4.87	19	2.38	2.35
4	5.67	5.63		12	4.58	4.51	20	2.17	2.09
5	7.21	7.33		13	4.20	4.05	21	1.97	1.99
6	5.72	5.78		14	3.80	3.86	22	1.79	1.78
7	5.83	5.64		15	3.46	3.40	23	1.63	1.63
8	6.43	6.45		16	3.16	3.21	24	1.48	1.42

N = 5 R = 50.000

	ber.	sim.		ber.	sim.		ber.	sim.	
1	0.00	0.00		9	5.17	5.03	17	3.01	2.97
2	0.00	0.00		10	4.71	4.83	18	2.81	2.81
3	0.47	0.52		11	4.39	4.43	19	2.62	2.65
4	3.16	3.33		12	4.21	4.17	20	2.44	2.37
5	5.04	5.11		13	3.98	3.91	21	2.28	2.18
6	4.18	4.12		14	3.70	3.75	22	2.12	2.16
7	4.38	4.22		15	3.45	3.68	23	1.98	1.93
8	5.22	5.39		16	3.23	3.20	24	1.85	1.89

N = 6 R = 60.000

	ber.	sim.		ber.	sim.		ber.	sim.	
1	0.00	0.00		9	4.44	4.34	17	2.96	3.04
2	0.00	0.00		10	4.14	4.10	18	2.80	2.83
3	0.14	0.13		11	3.91	4.00	19	2.64	2.66
4	1.75	1.75		12	3.82	3.70	20	2.50	2.50
5	3.62	3.54		13	3.68	3.67	21	2.36	2.25
6	3.11	2.99		14	3.48	3.53	22	2.23	2.16
7	3.35	3.54		15	3.29	3.43	23	2.11	2.15
8	4.27	4.13		16	3.12	3.13	24	1.99	2.01