Hydrologie und Flussgebietsmanagement

o.Univ.Prof. DI Dr. H.P. Nachtnebel

Institut für Wasserwirtschaft, Hydrologie und konstruktiver Wasserbau

Wiederholung Statistische Grundlagen

- Definitionen der statistischen Grundlagen
  - Grundgesamtheit / Stichprobe / Wahrscheinlichkeit
  - Absolute / relative Häufigkeit
  - Histogramm / Dichte- / Verteilungsfunktion
  - Summenlinie / Dauerlinie
- Verteilungen
  - Parameter zur Beschreibung
  - Normalverteilung
  - Standardisierung
- Begriffe
  - Jährlichkeit
  - Wiederkehrintervall

Empirische Verteilung	**Theoretische Verteilung**
Stichprobe, endlich	Grundgesamtheit, unendlich
Häufigkeitsverteilung (Histogramm)	Dichtefunktion f(x)
Summenhäufigkeit	Verteilungsfunktion F(x)
Relative Häufigkeit	Wahrscheinlichkeit
Empirische Wahrscheinlichkeiten = Häufigkeiten	
Mittelwert … $\bar{x}$	Mittelwert … μ
Standardabweichung … s	Standardabweichung … δ

Generelle Vorgangsweise bei Extremwertanalyse

- Auswahl der Stichprobe
- Auswahl einer Verteilung
- Anpassung der Verteilung an Stichprobe
- Ermittlung von Schätzwerten (Bemessungsgrößen)
- Schätzung der Unsicherheit der Aussage

Was ist ein Extremwert / Q_T

- Extremwert
  - Etwas, das selten auftritt
    - Hochwasser
    - Niederwasser
    - Lange niederschlagsfreie Perioden
    - Hagel, Schnee (saisonales Auftreten)
- Ermittlung von Q_T (Bemessungsgröße)
    - Quantil Q_T ist gesuchtes Extremereignis mit Wiederkehrintervall T bzw. Wahrscheinlichkeit 1/T
    - Ermittlung:
      - 1. Berechnung
      - 2. Grafisch

Beispiele

- Wiederkehrinterval T= 100 Jahre bedeutet, dass ein Ereignis im Durchschnitt (Mittel) alle hundert Jahre einmal auftritt
- Die Auftrittswahrscheinlichkeit W_A für ein Jahr ist daher 1/T = 0.01

Wie groß ist die Wahrscheinlichkeit für ein HQ_{100} in hundert Jahren ?

Beispiele

- Wiederkehrinterval T= 100 Jahre bedeutet, dass ein Ereignis im Durchschnitt (Mittel) alle hundert Jahre einmal auftritt
- Die Auftrittswahrscheinlichkeit W_A für ein Jahr ist daher 1/T = 0.01

Wie groß ist die Wahrscheinlichkeit für ein HQ_{100} in hundert Jahren ?

$W_{A,1}=1/100$, $W_{N,1}=1-1/100$

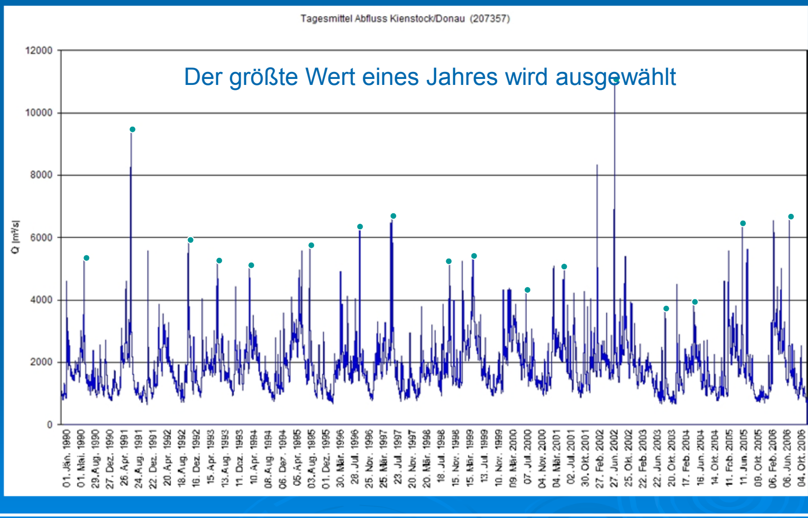
$W_{N,2}=W_{N,1}*W_{N,1}= (1-1/100)*(1-1/100)$

$W_{N,100}=W_{N,1}^{100} =(1-1/100)^{100}$

$W_{A,100}=1-W_{N,100} = 1-(1-1/100)^{100}= 63{,}4\ \%$

- Für derartige Fragen Anwendung des Binomialsatzes!!

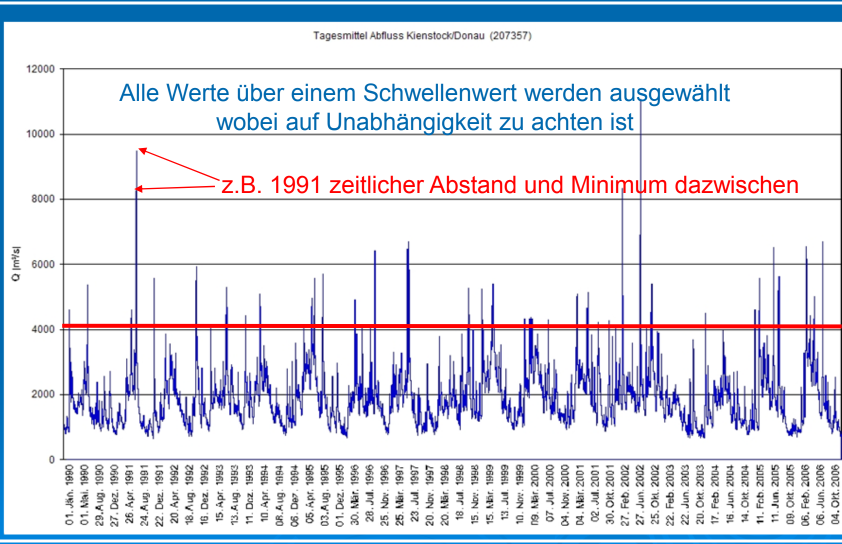
Jahresreihe



Extremwertstatistik

Seite 7

Partielle Reihe



Extremwertstatistik

Seite 8

Hinweise

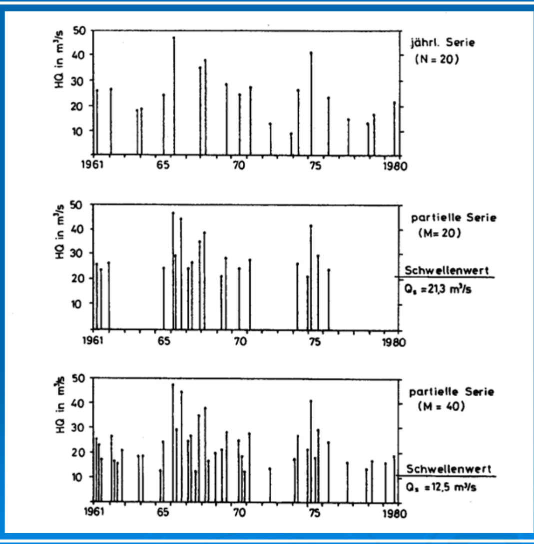
- Frage: wo liegt Schwellenwert?

  Faustformel: 3x soviel Werte als Beobachtungsjahre in der Stichprobe

- Wann wählt man Jahresreihe oder partielle Reihe ?

  Bei kurzen Beobachtungsreihen partielle Reihe
  Bei saisonaler Auswertung auch

Vergleich Jahresreihe mit partieller Reihe



Auswahl einer Verteilung

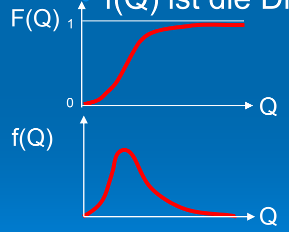
- Log-Normalverteilung
- Gumbelverteilung
- Log-Gumbelverteilung
- Pearson III-Verteilung
- Log-Pearson III Verteilung
- Weibull Verteilung
- Wakeby Verteilung
- Gamma Verteilung
- ….

Quantil und Verteilung

- Zusammenhang zwischen Q und $P(Q>Q_T)$

$$P(Q \le Q_T) = 1 - 1/T = F(Q) \qquad P(Q > Q_T) = 1 - P(Q \le Q_T)$$

- F(Q) ist gewählte Verteilungsfunktion
- f(Q) ist die Dichtefunktion



Gumbelverteilung

- Gewählte Verteilungsfunktion: Gumbel
  - 2parametrig (Schätzung)
  - Doppelt exponentiell
  - Linksseitig mit 0 begrenzt, rechtsseitig unbegrenzt

$$F(x_T) = e^{-e^{\frac{-a+x_T}{c}}} = 1 - \frac{1}{T} \quad \Rightarrow \log arithmieren$$

$$-e^{\frac{-a+x_T}{c}} = \ln\left(1-\frac{1}{T}\right) \quad \Rightarrow \log arithmieren \quad \Big|*(-1)$$

$$\frac{-a+x_T}{c} = \ln\left\{-\ln\left(1-\frac{1}{T}\right)\right\} \qquad \Rightarrow \frac{a}{c} + \frac{1}{c} = x_T$$

➔ Geradengleichung bei 2x logarithmieren der Gumbelverteilung

Grafische Ermittlung von QT

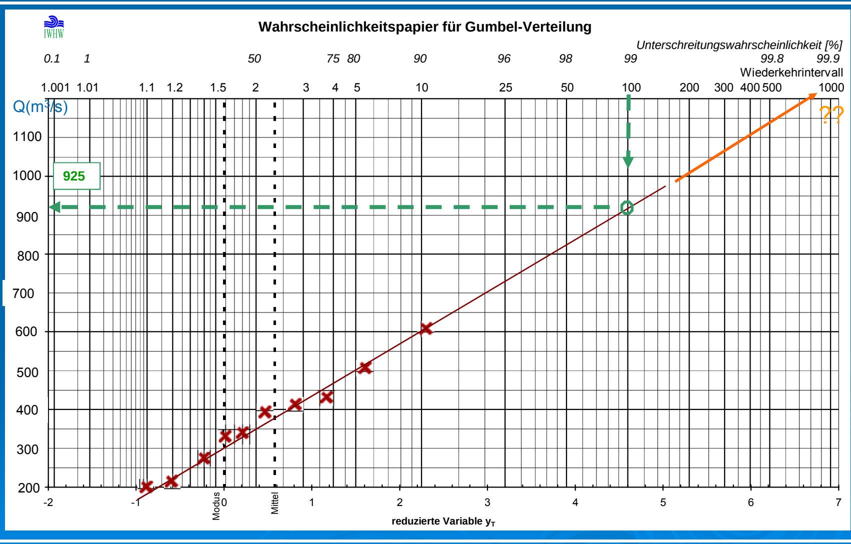
Jahr	Qmax	Rang	T	Weib.
1950	342	6	1,7	1,8
1951	415	4	2,5	2,8
1952	199	10	1	1,1
1953	278	8	1,3	1,4
1954	512	2	5	5,5
1955	333	7	1,4	1,6
1956	395	5	2	2,2
1957	607	1	10	11
1958	212	9	1,1	1,2
1959	437	3	3,3	3,7

Festlegung der Jährlichkeit T (Wiederkehrintervall)

- Beachten:
- 2 gleiche Werte
- Partielle Reihen
- Plotting Positions
  - Korrektur mittels Weibull

$$F(x) = \frac{k}{n+1}$$

Grafische Ermittlung von Q_T



Allgemeine Aussagen

1. T soll nicht größer sein als die 3-fache Beobachtungsdauer
2. Erwartungswert und Unsicherheit angeben

Rechnerische Ermittlung von Q_T bzw. x_T

- Schätzwerte

$$F(x_T) = e^{-e^{\frac{-a+x_T}{c}}} = 1 - \frac{1}{T}$$

- Parameter
  a (Maßstabsparameter)
  c (Lageparameter)

$$a = \bar{x} - 0{,}5772 * c$$

$$c = \frac{\pi}{\sqrt{6}s_x}$$

- Hydrologische Grundgleichung

$$x_T = \bar{x} + K(T) * s_x$$

- Entsprechung bei NV

$$x_T = \bar{x} + u(T) * s_x$$

Rechnerische Ermittlung von Q_T

Häufigkeitsfaktor K_T

n	Wiederholungszeitspanne in Jahren T								
	1.053	1.111	1.25	2	5	10	20	50	100
5	-1.963	-1.631	-1.179	-0.116	1.313	2.26	3.168	4.343	5.224
10	-1.677	-1.4	-1.023	-0.136	1.058	1.848	2.606	3.587	4.323
15	-1.578	-1.32	-0.969	-0.143	0.967	1.703	2.408	3.321	4.005
20	-1.525	-1.277	-0.94	-0.148	0.919	1.625	2.302	3.179	3.836
25	-1.492	-1.251	-0.922	-0.151	0.888	1.575	2.235	3.089	3.728
30	-1.468	-1.232	-0.91	-0.153	0.866	1.541	2.188	3.026	3.653
35	-1.451	-1.218	-0.901	-0.154	0.85	1.515	2.153	2.979	3.598
40	-1.438	-1.207	-0.893	-0.155	0.838	1.495	2.126	2.943	3.554
45	-1.427	-1.198	-0.887	-0.156	0.828	1.479	2.104	2.913	3.519
50	-1.418	-1.191	-0.883	-0.157	0.82	1.466	2.086	2.889	3.491
55	-1.41	-1.185	-0.879	-0.157	0.813	1.455	2.071	2.869	3.467
60	-1.404	-1.18	-0.875	-0.158	0.807	1.446	2.059	2.852	3.446
65	-1.398	-1.176	-0.872	-0.158	0.802	1.438	2.047	2.837	3.428
70	-1.394	-1.172	-0.869	-0.159	0.797	1.43	2.038	2.824	3.413
75	-1.389	-1.168	-0.867	-0.159	0.793	1.424	2.029	2.812	3.399
80	-1.386	-1.165	-0.865	-0.159	0.79	1.419	2.021	2.802	3.387
85	-1.382	-1.162	-0.863	-0.16	0.787	1.413	2.015	2.793	3.376
90	-1.379	-1.16	-0.862	-0.16	0.784	1.409	2.008	2.784	3.366
95	-1.376	-1.158	-0.86	-0.16	0.781	1.405	2.003	2.777	3.357
100	-1.374	-1.155	-0.859	-0.16	0.779	1.401	1.998	2.77	3.349

$$y_T = \bar{y} + K(T) * s_y$$

$$y_T = -\ln \ln\left(\frac{T}{T-1}\right)$$

$$x_{100} = \bar{x} + K(T) * s_x =$$

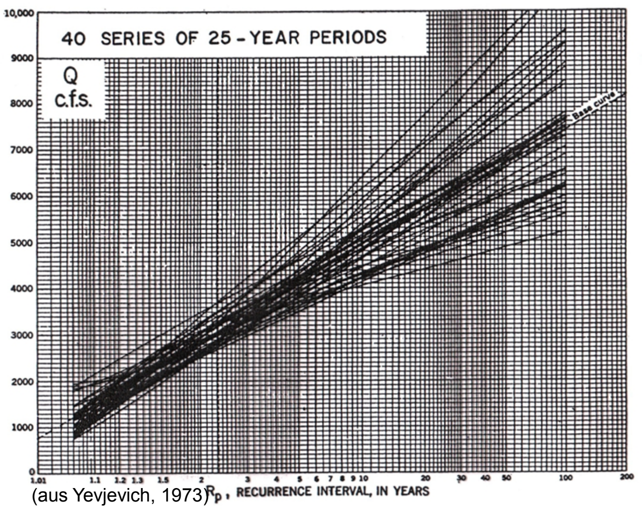
$$= 373 + 4{,}323 * 128{,}3$$

$$= 928\ m^3/s$$

Bemessungswert und Schätzfehler

- An
- An
- Sc ... länge
T

en



(aus Yevjevich, 1973)

Figure 4.3. Frequency distributions of 40 samples, each of a 25-year size, of the maximum flood discharges on a Gumbel paper graph. Each sample is drawn from a finite population of 1,000 values without replacement. These values follow the double exponential distribution. [According to Benson, 7]

Berücksichtigung des Schätzfehlers

- Annahme: die Messungen sind perfekt
  aber man weiß nicht d
- Annahme: das Modell ist
- Schätzfehler: nimmt ab m
  nimmt zu m
  ist NV verte

n	Wiederholungszeitspanne in Jahren								
	1.053	1.111	1.25	2	5	10	20	50	100
5	1.732	1.438	1.089	0.939	2.096	3.032	3.956	5.167	6.08
10	1.477	1.249	0.992	0.93	1.854	2.62	3.383	4.387	5.146
15	1.393	1.188	0.964	0.927	1.77	2.476	3.181	4.113	4.817
20	1.349	1.157	0.949	0.925	1.725	2.399	3.075	3.967	4.643
25	1.322	1.138	0.941	0.924	1.697	2.351	3.007	3.875	4.532
30	1.303	1.125	0.935	0.923	1.677	2.317	2.96	3.81	4.455
35	1.289	1.115	0.93	0.922	1.663	2.292	2.925	3.762	4.397
40	1.279	1.108	0.927	0.922	1.651	2.272	2.897	3.725	4.353
45	1.27	1.102	0.925	0.921	1.642	2.257	2.876	3.695	4.317
50	1.263	1.097	0.923	0.921	1.635	2.244	2.858	3.671	4.287
55	1.257	1.093	0.921	0.921	1.629	2.233	2.843	3.65	4.263
60	1.252	1.089	0.919	0.921	1.624	2.224	2.83	3.633	4.242
65	1.248	1.086	0.918	0.92	1.619	2.216	2.819	3.617	4.223
70	1.244	1.084	0.917	0.92	1.615	2.209	2.809	3.604	4.207
75	1.241	1.082	0.916	0.92	1.611	2.203	2.8	3.592	4.193
80	1.238	1.08	0.915	0.92	1.608	2.198	2.793	3.582	4.181
85	1.235	1.078	0.914	0.92	1.605	2.193	2.786	3.572	4.169
90	1.233	1.076	0.914	0.92	1.603	2.188	2.78	3.564	4.159
95	1.231	1.075	0.913	0.92	1.601	2.184	2.774	3.556	4.15
100	1.229	1.073	0.912	0.92	1.599	2.181	2.769	3.549	4.141

Tabelle δ_T

$$x_T \pm u(\alpha) * s_T = x_T \pm u(\alpha) * \delta_T \frac{s_x}{\sqrt{n}}$$

$$\delta_T = \sqrt{1 + 1{,}14 * K_T + 1{,}1 * K_T^2}$$

K_T, δ_T und $u(\alpha)$ aus Tabellen

$$\delta_T = \sqrt{1 + 1{,}14 * 4{,}323 + 1{,}1 * 4{,}323^2} = 5{,}146$$

$$928 \pm 1{,}960 \cdot 5{,}146 \cdot 128{,}3 / \sqrt{10} = 928\ m^3/s \pm 409\ m^3/s$$

P	0.975
z	1.960

$u(\alpha)$

δ_T – Wert / Red. Zufallsvariable der Gumbel VT

$$\delta_T = \sqrt{1 + 1.14 K_T + 1.1 K_T^2}$$

Stich-proben-umfang n	Wiederholungszeitspanne in Jahren 1.053	1.111	1.25	2	5	10	20	50	100
5	1.732	1.438	1.089	0.939	2.096	3.032	3.956	5.167	6.08
10	1.477	1.249	0.992	0.93	1.854	2.62	3.383	4.387	5.146
15	1.393	1.188	0.964	0.927	1.77	2.476	3.181	4.113	4.817
20	1.349	1.157	0.949	0.925	1.725	2.399	3.075	3.967	4.643
25	1.322	1.138	0.941	0.924	1.697	2.351	3.007	3.875	4.532
30	1.303	1.125	0.935	0.923	1.677	2.317	2.96	3.81	4.455
35	1.289	1.115	0.93	0.922	1.663	2.292	2.925	3.762	4.397
40	1.279	1.108	0.927	0.922	1.651	2.272	2.897	3.725	4.353
45	1.27	1.102	0.925	0.921	1.642	2.257	2.876	3.695	4.317
50	1.263	1.097	0.923	0.921	1.635	2.244	2.858	3.671	4.287
100	1.229	1.073	0.912	0.92	1.599	2.181	2.769	3.549	4.141

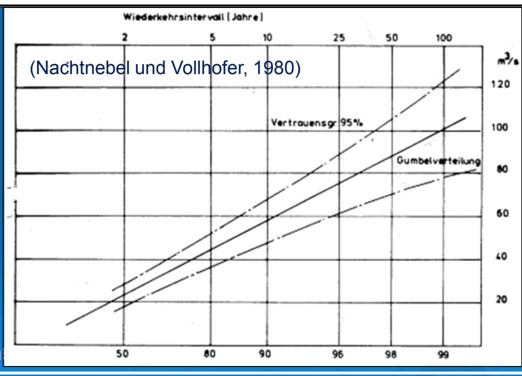
δ_T - Wert der Gumbel-Verteilung in Abhängigkeit von Jährlichkeit und Stichprobenumfang

T	1.01010101	2	5	10	25	50	100	200	500	1000
$P_ü$	99	50	20	10	4	2	1	0.5	0.2	0.1
y_T	-1.5272	0.3665	1.4999	2.2504	3.1985	3.9019	4.6001	5.2958	6.2136	6.9073

Reduzierte Zufallsvariable y_T = -ln ln (T_R/T_R-1) als Funktion des Wiederkehrintervalls (Gumbelverteilung)

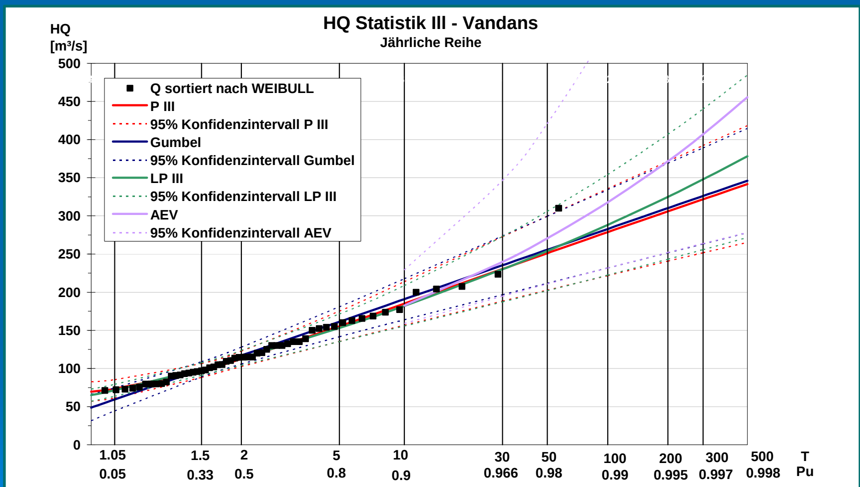
Schätzwert und Schätzfehler

Abflussgeschehen		Maxima					
Bad Schallerbach							
Wiederkehrintervall (Jahre)		2	5	10	25	50	100
Verteilung				HQ m³/s			
Gumbel		23,46	43,74	57,44	74,57	87,19	99,81
Vertrauensbereich	+/-	4,32	7,85	11,20	14,90	18,45	21,62
Log Pearson III		20,87	35,95	48,95	64,14	87,56	108,90



http://www.uni-kon

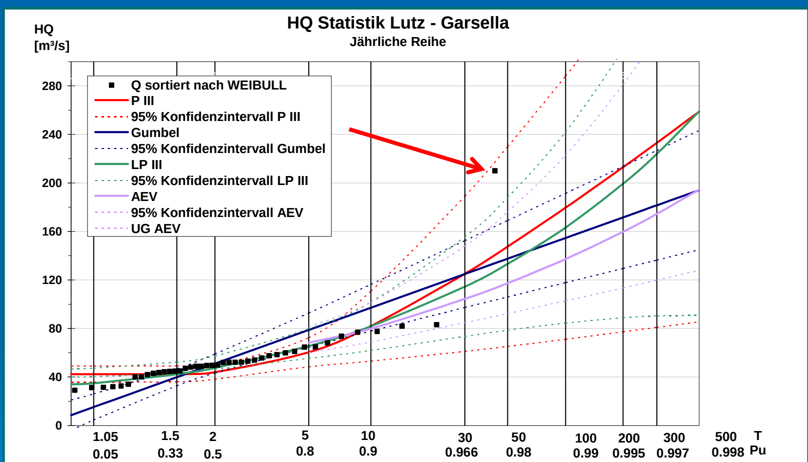
Vergleich verschiedener Verteilungen



(Nachtnebel und Stanzel, 2007)

Extremwertstatistik Seite 23

Ausreisser ?

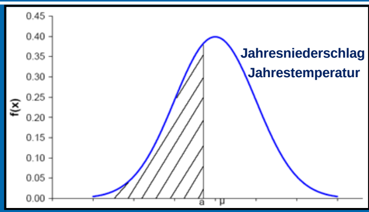


(Nachtnebel und Stanzel, 2007)

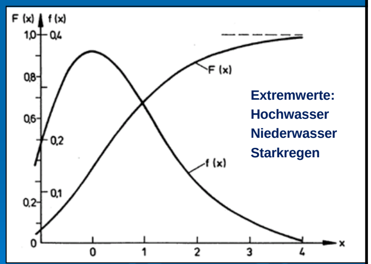
Extremwertstatistik Seite 24

Verteilungen Übersicht 1

- Normalverteilung
  - 2-parametrig
  - Symmetrisch
  - Beidseitig unbegrenzt

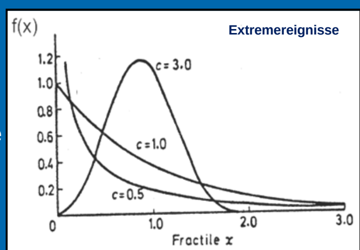


- Gumbel-Verteilung
  - 2parametrig
  - Doppelt exponentiell $F(x)=e^{-e^{\frac{a+x}{c}}}$
  - Asymmetrisch mit fester Schiefe $c_s = 1{,}1396$
  - Parameter a – Lageparameter, Modalwert $a = \bar{x} - 0{,}5772 * c$
  - Parameter c - Maßstabsparameter
  - Rechtsseitig unbegrenzt $c = \frac{1{,}28255}{s_x}$



Verteilungen Übersicht 2

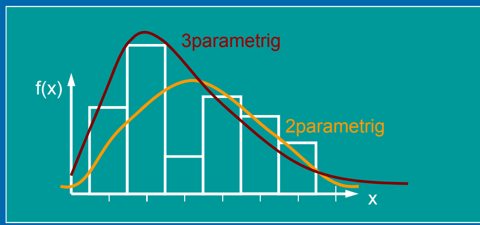
- Weibull-Verteilung
  - Sonderfall der Kritsky-Menkel Verteilung
  - 3-parametrig
  - Asymmetrisch ohne fester Schiefe
  - Rechtsseitig unbegrenzt



- Welche Verteilung ist nun die beste zur Berechnung eines Hochwasserereignisses?
  - Keine eindeutige Aussage – aber Empfehlung:
    - Gumbel
    - Weibull
    - Pearson
    - GammaVerteilung

Ergebnisinterpretation

- Verschiedene Verteilungen – verschiedene Ergebnisse
  - 2parametrig vs. 3parametrig
  - Auswirkung der Verteilungswahl auf Kriterium:
    - Gumbel liefert für kleine Jährlichkeiten höhere Werte
    - Pearson liefert dies für große Jährlichkeiten



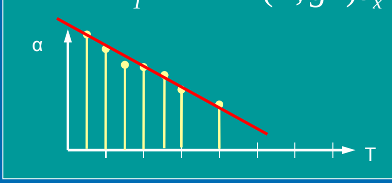
<u>IMMER:</u>

Angabe des Ergebnisses mit
- Erwartungswert
- und Konfidenzintervall

Niederwasserstatistik

- Anwendung:
  - Wasserentnahme
  - Einleitung in Vorfluter – Beispiel: thermische Belastung

$$x_T = \bar{x} + K(T, g^*)s_x$$



- KEIN negativer Abfluss
- Für Österreich:
  - Log Pearson
  - Log Gumbel

Zusammenfassung Extremwertstatistik

- Jahresreihe vs. partielle Reihe
- Grafische und rechnerische Ermittlung eines Extremereignisses
  - Grafisch – Wahrscheinlichkeitspapier
  - Rechnerisch
- Hydrologische Grundgleichung
- Angabe des Ergebnisses mit
  - Erwartungswert
  - und Konfidenzintervall
- Überblick über Verteilungen
- Anwendung der Verteilungen