Monte Carlo Methoden

vfvf-Mittagsseminar
Konzepte der Versicherungsmathematik einfach erklärt

Universität Oldenburg,18. Juni 2024

Dietmar Pfeifer

Schwerpunkt Versicherungs- und Finanzmathematik

CARL VON OSSIETZKY universität OLDENBURG
FAKULTÄT V
INSTITUT FÜR MATHEMATIK

Agenda

1. Zufallszahlen
2. Erzeugung spezieller statistischer Verteilungen
3. Ein Anwendungsbeispiel aus der Versicherungswirtschaft
4. Literatur

1. Zufallszahlen

1. Zufallszahlen

Praktisch alle computer-orientierten Verfahren zur Simulation (Monte Carlo Verfahren) beruhen auf so genannten *Pseudo-Zufallszahlen* z_n, die algorithmisch erzeugt werden und die die Eigenschaften stochastisch unabhängiger, über dem Intervall $[0,1]$ stetig gleichverteilter Zufallsvariablen (so genannte *Standard-Zufallszahlen*) imitieren. Ein klassischer Ansatz stammt aus der Zahlentheorie (Algebra): Hier wählt man

$$\boxed{z_n := \frac{u_n}{m},\ u_{n+1} := a \cdot u_n \mod m}$$

mit geeigneten natürlichen Zahlen $a, m \in \mathbb{N}$ und einem von Null verschiedenen Startwert $0 < u_0 < m$. Die Rechenvorschrift $u_{n+1} := a \cdot u_n \mod m$ bedeutet dabei, dass man den zwischen 0 und $m-1$ liegenden Divisionsrest von $a \cdot u_n$ bei Teilung durch m ermittelt.

Beispiel 1: $m = 45659$ (Primzahl), $a = 1113$, $u_0 = 5555$

$$z_0 = \frac{u_0}{m} = 0{,}12166\ldots;\ \frac{a \cdot u_0}{m} = \frac{6.182.715}{45.659} = 135.41065\ldots; z_1 = 0{,}41065\ldots$$

n	0	1	2	3	4	5	6	7	8	9	10
z_n	0,122	0,411	0,057	0,062	0,595	0,306	0,099	0,084	0,069	0,712	0,890

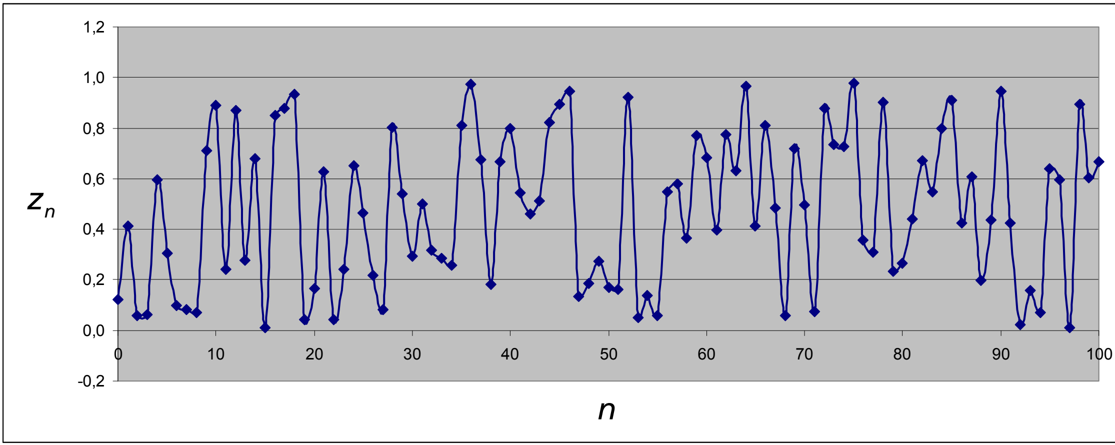
n	11	12	13	14	15	16	17	18	19	20
z_n	0,241	0,872	0,276	0,678	0,009	0,848	0,877	0,932	0,043	0,168

n	21	22	23	24	25	26	27	28	29	30
z_n	0,625	0,042	0,241	0,652	0,465	0,218	0,082	0,802	0,540	0,292

n	31	32	33	34	35	36	37	38	39	40
z_n	0,498	0,318	0,286	0,258	0,812	0,973	0,676	0,181	0,666	0,797

n	41	42	43	44	45	46	47	48	49	50
z_n	0,543	0,460	0,513	0,821	0,894	0,945	0,132	0,184	0,273	0,168

Beispiel 1: $m = 45659$ (Primzahl), $a = 1113$, $u_0 = 5555$

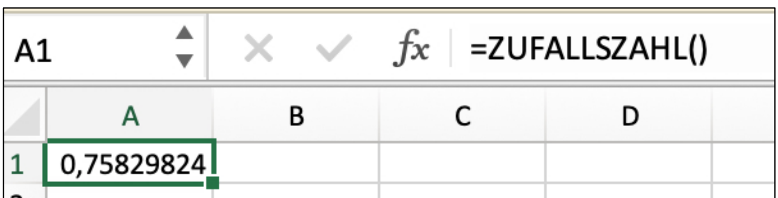


Um zu vermeiden, dass der Algorithmus die Zahl 0 erzeugt, wird für m in der Regel eine Primzahl oder Primzahlpotenz verwendet; a und m müssen dabei teilerfremd sein.

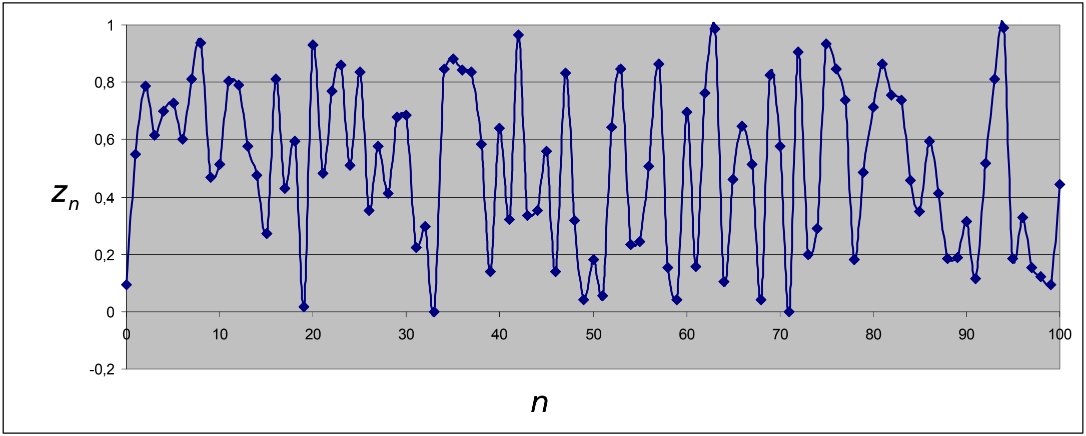
Leider können vor allem im Mehrdimensionalen durch solch einfache Verfahren unerwünschte Nebeneffekte auftreten (z.B. Konzentration konsekutiver Zufallszahlen auf niedrig-dimensionale Gitter). In modernen Ansätzen werden daher Matrix-Multiplikationen verwendet, um unmittelbar höherdimensionale Zufallsvektoren zu erzeugen, oder mehrere rekursive Verfahren, die durch eine geeignete Schieberegister-Technik miteinander verschnitten werden.

Für praktisch alle Programmiersprachen und Tabellenkalkulationssysteme gibt es Befehle zur Erzeugung von Standard-Zufallszahlen.

In EXCEL ist das etwa: f_x = ZUFALLSZAHL()



Beispiel 2:



2. Erzeugung spezieller statistischer Verteilungen

2. Erzeugung spezieller statistischer Verteilungen

Die so genannte Inversionsmethode beruht auf folgender allgemeingültiger Beobachtung: bezeichnet man wie üblich mit F die Verteilungsfunktion einer reellen (diskreten, stetigen oder gemischt stetig-diskreten) Zufallsvariablen X und die *Quantilfunktion* F^{-1} durch

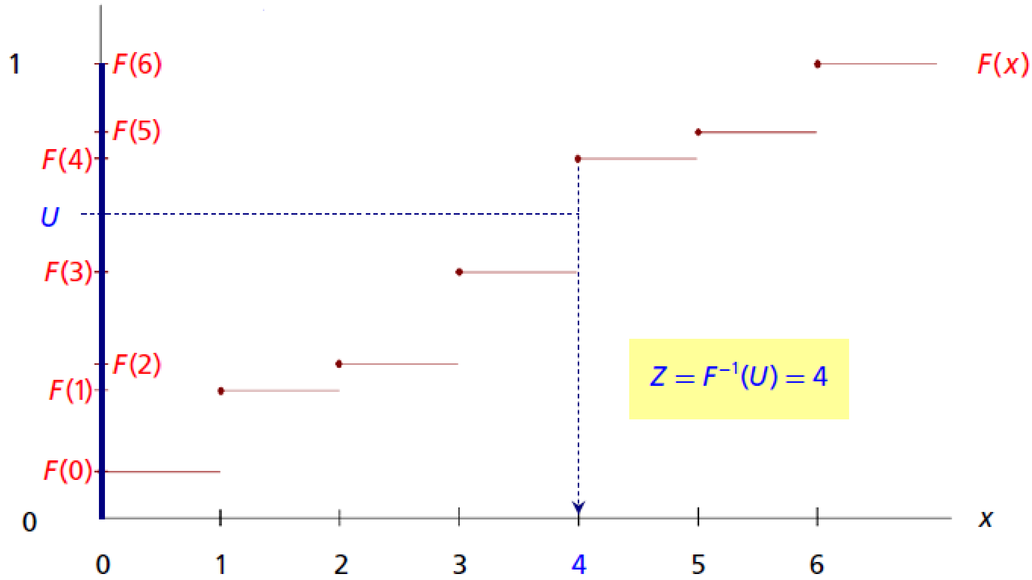
$$\boxed{F^{-1}(u) := \inf\{x \in \mathbb{R} \mid F(x) \geq u\},\ 0 < u < 1,}$$

so gilt:

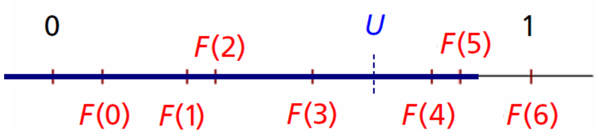
- $U := F(X)$ ist eine Standard-Zufallszahl, wenn F *stetig* ist;
- $Z := F^{-1}(U)$ ist – ohne Einschränkung – wie X verteilt, wenn U eine Standard-Zufallszahl ist.

Die Inversionsmethode ist insbesondere für die Erzeugung diskreter Verteilungen geeignet. Sie kann unter anderem in EXCEL leicht implementiert werden.

Eine Veranschaulichung des Verfahrens zeigt die folgende Graphik für eine diskrete Zufallsvariable mit den Werten $\{0,1,\cdots,6\}$.


1
F(6)
F(5)
F(4)
U
F(3)
F(2)
F(1)
F(0)
0
F(x)
x
0
1
2
3
4
5
6
Z = F^{-1}(U) = 4

Nach Drehung der Ordinatenachse:



Hier wird durch U der diskrete Zufallswert „4“ generiert. Fällt die Standard-Zufallszahl U in das Intervall $\left(F(k-1), F(k)\right]$ mit $F(-1)=0$, so wird $Z := k$ gesetzt; dann gilt nämlich

$$\boxed{P(Z=k)=P\big(F(k-1)<U\leq F(k)\big)=F(k)-F(k-1).}$$

Beispiel 4: diskrete Gleichverteilung über der Menge $\{1,2,\cdots,n\}$: hier ist

$$F(k)=\frac{k}{n},\ k=0,\cdots,n$$

mit $Z=k$ genau dann, wenn $\frac{k-1}{n}<U\leq\frac{k}{n}$ ist. Damit erhält man einfacher

$$Z=\lceil n\cdot U\rceil \quad \text{oder} \quad Z=\lfloor n\cdot U\rfloor+1,$$

wobei

$\lceil z\rceil:=\min\{m\in\mathbb{Z}\,|\,z\leq m\}$ die nach oben aufgerundete

$\lfloor z\rfloor:=\max\{m\in\mathbb{Z}\,|\,z\geq m\}$ die nach unten abgerundete

nächste ganze Zahl zu z ist.

Mit der Inversionsmethode lassen sich darüber hinaus alle stetigen Verteilungen einfach generieren, die eine geschlossene Darstellung der Verteilungsfunktion F und ihrer Umkehrfunktion (Quantilfunktion) $Q = F^{-1}$ besitzen. Dazu gehören u.a. die Exponential-Verteilung, Weibull-Verteilung, Fréchet-Verteilung, Pareto-Verteilung, Logistische und Loglogistische Verteilung.

In EXCEL existieren dazu u.a. folgende weitere Funktionen:

Normalverteilung:

$F(x): f_x = \text{normvert}(x,\mu,\sigma,\text{wahr}),\ x \in \mathbb{R}\ \ (\mu \in \mathbb{R},\ \sigma > 0)$

$Q(u): f_x = \text{norminv}(u,\mu,\sigma),\ 0 < u < 1.$

Syntax

NORMVERT(**x**;**Mittelwert**;**Standabwn**;**Kumuliert**)

x ist der Wert der Verteilung (Quantil), dessen Wahrscheinlichkeit Sie berechnen möchten.

Mittelwert ist das arithmetische Mittel der Verteilung.

Standabwn ist die Standardabweichung der Verteilung.

Kumuliert ist der Wahrheitswert, der den Typ der Funktion bestimmt. Ist **Kumuliert** mit WAHR belegt, gibt NORMVERT den Wert der Verteilungsfunktion (kumulierte Dichtefunktion) zurück. Ist **Kumuliert** mit FALSCH belegt, gibt NORMVERT den Wert der Dichtefunktion zurück.

Hinweise

- Ist **Mittelwert** oder **Standabwn** kein numerischer Ausdruck, gibt NORMVERT den Fehlerwert #WERT! zurück.
- Ist **Standabwn** ≤ 0, gibt NORMVERT den Fehlerwert #ZAHL! zurück.
- Ist **Mittelwert** = 0, **Standabwn** = 1 und **Kumuliert** = TRUE, gibt NORMVERT die Standardnormalverteilung (STANDNORMVERT) zurück.
- Die Gleichung der Dichtefunktion der Normalverteilung (**Kumuliert** = FALSE) lautet:

$$f(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\left(\frac{(x-\mu)^2}{2\sigma^2}\right)}$$

- Wenn **Kumuliert** = TRUE, gibt die Formel das Integral einer gegebenen Formel von der negativen Unendlichkeit bis x zurück.

Gammaverteilung:

$F(x): f_x = \text{gammavert}(x, \alpha, \beta, \text{wahr}),\ x \in \mathbb{R}\ \ (\alpha, \beta > 0)$

$Q(u): f_x = \text{gammainv}(u, \alpha, \beta), 0 < u < 1.$

Hierbei ist $\beta = \dfrac{1}{\lambda}$ in der Standardnotation der Gammaverteilung.

Syntax

GAMMAVERT(**x**;**Alpha**;**Beta**;**Kumuliert**)

x ist der Wert (Quantil), dessen Wahrscheinlichkeit (1-Alpha) Sie berechnen wollen.

Alpha ist ein Parameter der Verteilung.

Beta ist ein Parameter der Verteilung. Wenn *Beta* = 1, gibt **GAMMAVERT** die Standard-Gammaverteilung zurück.

Kumuliert ist der Wahrheitswert, der den Typ der Funktion bestimmt. Ist *Kumuliert* mit **WAHR** belegt, berechnet **GAMMAVERT** den Wert der Verteilungsfunktion, also die Wahrscheinlichkeit, dass die Anzahl zufällig auftretender Ereignisse zwischen 0 und einschließlich *x* liegt. Ist *Kumuliert* mit **FALSCH** belegt, liefert **GAMMAVERT** den Wert der Dichtefunktion.

Hinweise

- Ist *x*, *Alpha* oder *Beta* kein numerischer Ausdruck, gibt **GAMMAVERT** den Fehlerwert **#WERT!** zurück.
- Ist *x* < 0, gibt **GAMMAVERT** den Fehlerwert **#ZAHL!** zurück.
- Ist *Alpha* ≤ 0 oder *Beta* ≤ 0, gibt **GAMMAVERT** den Fehlerwert **#ZAHL!** zurück.
- Die Gleichung für die Dichtefunktion der Gammaverteilung lautet:

$$f(x;\alpha,\beta)=\frac{1}{\beta^{\alpha}\Gamma(\alpha)}x^{\alpha-1}e^{-\frac{x}{\beta}}$$

Die Dichtefunktion der standardisierten Gammaverteilung lautet:

$$f(x;\alpha)=\frac{x^{\alpha-1}e^{-x}}{\Gamma(\alpha)}$$

- Ist *Alpha* = 1, gibt **GAMMAVERT** die Exponentialverteilung mit Folgendem zurück:

$$\lambda=\frac{1}{\beta}$$

- Ist *Alpha* eine positive ganze Zahl, wird **GAMMAVERT** auch als "Erlang-Verteilung" bezeichnet.

Betaverteilung:

$$F(x): f_x = \text{betavert}(x,\alpha,\beta,[A],[B]),\ 0 < x < 1\ \ (\alpha,\beta > 0,\ 0 < A < B)$$

$$Q(u): f_x = \text{betainv}(u,\alpha,\beta,[A],[B]),\ 0 < u < 1.$$

Syntax

BETAVERT(**X;Alpha;Beta**;A;B)

X ist der Wert, an dem die Funktion im Intervall zwischen **A** und **B** ausgewertet werden soll.

Alpha ist ein Parameter der Verteilung.

Beta ist ein Parameter der Verteilung.

A ist die optionale untere Grenze des Intervalls für **X**.

B ist die optionale obere Grenze des Intervalls für **X**.

Hinweise

- Ist eines der Argumente nichtnumerisch, gibt **BETAVERT** den Fehlerwert #WERT! zurück.
- Ist **Alpha** ≤ 0 oder **Beta** ≤ 0, gibt **BETAVERT** den Fehlerwert #ZAHL! zurück.
- Ist **X** < **A**, **X** > **B**, oder **A** = **B**, gibt **BETAVERT** den Fehlerwert #ZAHL! zurück.
- Wird für **A** und **B** kein Wert angegeben, verwendet **BETAVERT** die Standardverteilung, d. h. **A** = 0 und **B** = 1.

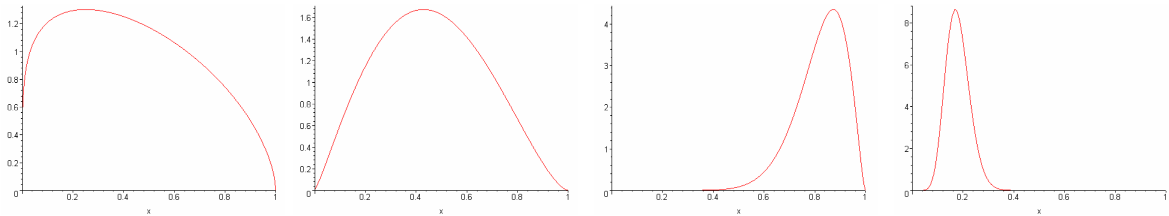
Die Standard-Betaverteilung mit den Parametern α und β besitzt die Verteilungsdichte

$$\boxed{f(x;\alpha,\beta)=\frac{x^{\alpha-1}(1-x)^{\beta-1}}{\mathcal{B}(\alpha,\beta)},\ 0<x<1.}$$

Hierbei bezeichnet $\mathcal{B}(\alpha,\beta)$ die sogenannte Beta-Funktion, die definiert ist durch

$$\boxed{\mathcal{B}(\alpha,\beta)=\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}}$$

mit der Euler'schen Gamma-Funktion $\Gamma(x)=\int\limits_0^\infty t^{x-1}e^{-t}\,dt$ für $x>0$.



graphische Darstellung von Dichten verschiedener Beta-Verteilungen

$\alpha = 1{,}2 \quad \beta = 1{,}6 \qquad \alpha = 2{,}2 \quad \beta = 2{,}6 \qquad \alpha = 12 \quad \beta = 2{,}6 \qquad \alpha = 12 \quad \beta = 55$

3. Ein Anwendungsbeispiel aus der Versicherungswirtschaft

3. Ein Anwendungsbeispiel aus der Versicherungswirtschaft

Wir simulieren den zeitlichen Verlauf einer Versicherungssparte über 10 Jahre in der Zukunft.

Eingabeparameter:

Wie viele Policen sind prozentual (zufällig) von einem Schaden betroffen? ($\longrightarrow$ Betaverteilung)

Wieviel Prozent der Versicherungssumme (zufällig) macht die Schadenhöhe aus? ($\longrightarrow$ Betaverteilung)

Wie hoch ist die (zufällige) jährliche Wachstumsrate der Policen? ($\longrightarrow$ Betaverteilung)

Wieviel Prozent der Versicherungssumme macht der Tarif aus? (fix)

Wie hoch ist der Kostensatz in Prozent der Prämien? (fix)

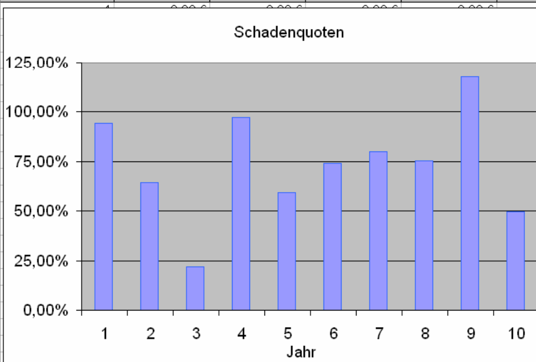
Ausgabe:

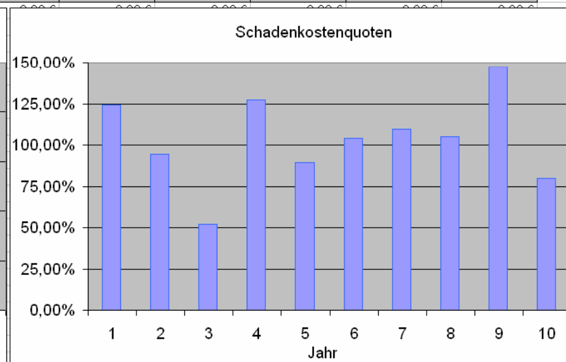
Jährliche Schadenquoten – Combined Ratios

Microsoft Excel - SimulationXLS

A1 | Prämie

	A	B	C	D	E	F	G	H	I	J	K	L
1	Prämie	247.572,61 €	256.395,68 €	274.248,62 €	299.901,12 €	323.873,07 €	380.986,60 €	386.133,28 €	407.025,47 €	440.557,76 €	488.303,35 €	
2	Bedarfsprämie	148.543,57 €	153.837,41 €	164.549,17 €	179.940,67 €	194.323,84 €	228.591,96 €	231.679,97 €	244.215,28 €	264.334,65 €	292.982,01 €	
3	N	51	33	12	59	47	60	64	62	101	47	
4	Betroffenheitsgrad	5,10%	3,18%	1,09%	4,88%	3,62%	3,92%	4,12%	3,80%	5,71%	2,40%	
5	Schaden	232.933,07 €	164.673,69 €	59.902,95 €	292.062,60 €	192.689,61 €	281.994,42 €	308.042,27 €	306.365,87 €	518.499,82 €	243.044,61 €	
6	Schadenquote	94,09%	64,23%	21,84%	97,39%	59,50%	74,02%	79,78%	75,27%	117,69%	49,77%	
7	Kosten	74.271,78 €	76.918,70 €	82.274,59 €	89.970,33 €	97.161,92 €	114.295,98 €	115.839,99 €	122.107,64 €	132.167,33 €	146.491,01 €	
8	Aufwand	307.204,85 €	241.592,40 €	142.177,53 €	382.032,93 €	289.851,53 €	396.290,40 €	423.882,26 €	428.473,51 €	650.667,14 €	389.535,61 €	
9	combined ratio	124,09%	94,23%	51,84%	127,39%	89,50%	104,02%	109,78%	105,27%	147,69%	79,77%	103,36%
10	Policenzahl	1.000	1.038	1.106	1.209	1.299	1.532	1.552	1.632	1.768	1.958	
11	Jahr	1	2	3	4	5	6	7	8	9	10	
34	23	0,00 €	0,00 €	0,00 €	0,00 €	0,00 €	0,00 €	0,00 €	0,00 €	0,00 €	0,00 €	

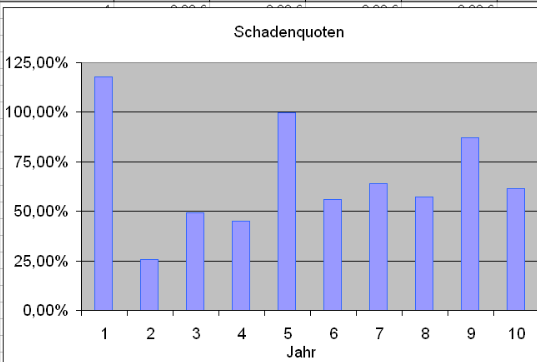


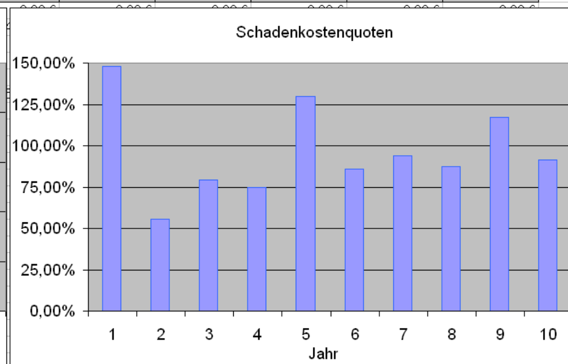


P / Z / Anz / Simulation / Q-Q-Plot

Microsoft Excel - SimulationXLS

	A	B	C	D	E	F	G	H	I	J	K	L
1	Prämie	247.572,61 €	264.501,45 €	285.201,98 €	302.567,36 €	327.278,91 €	375.658,06 €	433.956,97 €	507.145,09 €	559.297,61 €	619.906,24 €	
2	Bedarfsprämie	148.543,57 €	158.700,87 €	171.121,19 €	181.540,42 €	196.367,35 €	225.394,84 €	260.374,18 €	304.287,05 €	335.578,57 €	371.943,75 €	
3	N	59	17	32	28	72	45	51	57	98	82	
4	Betroffenheitsgrad	5,90%	1,59%	2,78%	2,30%	5,48%	2,99%	2,93%	2,79%	4,35%	3,28%	
5	Schaden	291.643,38 €	67.672,70 €	140.662,77 €	136.014,44 €	326.630,97 €	210.136,12 €	276.908,14 €	290.676,34 €	487.512,43 €	380.791,93 €	
6	Schadenquote	117,80%	25,59%	49,32%	44,95%	99,80%	55,94%	63,81%	57,32%	87,17%	61,43%	
7	Kosten	74.271,78 €	79.350,43 €	85.560,59 €	90.770,21 €	98.183,67 €	112.697,42 €	130.187,09 €	152.143,53 €	167.789,28 €	185.971,87 €	
8	Aufwand	365.915,16 €	147.023,13 €	226.223,36 €	226.784,65 €	424.814,65 €	322.833,54 €	407.095,24 €	442.819,86 €	655.301,71 €	566.763,81 €	
9	combined ratio	147,80%	55,59%	79,32%	74,95%	129,80%	85,94%	93,81%	87,32%	117,17%	91,43%	96,31%
10	Policenzahl	1.000	1.070	1.151	1.219	1.314	1.507	1.739	2.042	2.253	2.501	
11	Jahr	1	2	3	4	5	6	7	8	9	10	
34	23	0,00 €	0,00 €	0,00 €	0,00 €	0,00 €	0,00 €	0,00 €	0,00 €	0,00 €	0,00 €	



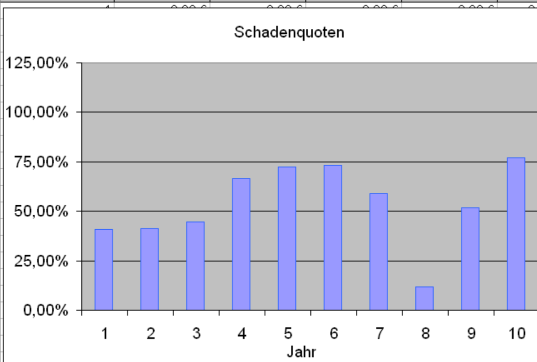


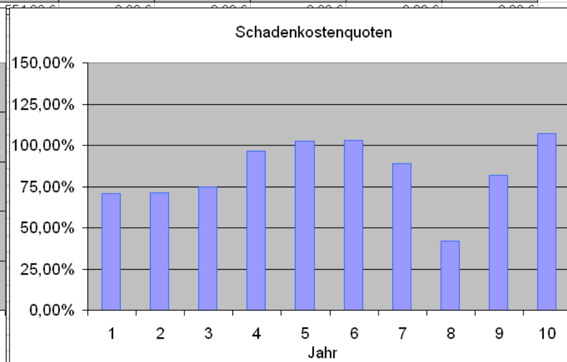
P / Z / Anz / Simulation / Q-Q-Plot

Microsoft Excel - SimulationXLS

A1 Prämie

	A	B	C	D	E	F	G	H	I	J	K	L
1	Prämie	247.572,61 €	263.921,76 €	282.894,17 €	328.182,43 €	345.212,92 €	363.957,08 €	427.688,21 €	492.105,33 €	494.168,32 €	558.871,77 €	
2	Bedarfsprämie	148.543,57 €	158.353,05 €	169.736,50 €	196.909,46 €	207.127,75 €	218.374,25 €	256.612,93 €	295.263,20 €	296.500,99 €	335.323,06 €	
3	N	21	23	31	44	46	53	46	12	53	87	
4	Betroffenheitsgrad	2,10%	2,15%	2,72%	3,34%	3,32%	3,63%	2,68%	0,61%	2,67%	3,86%	
5	Schaden	100.608,85 €	108.743,28 €	125.940,44 €	218.805,12 €	249.701,73 €	266.682,01 €	251.290,38 €	59.014,96 €	256.814,15 €	429.569,68 €	
6	Schadenquote	40,64%	41,20%	44,52%	66,67%	72,33%	73,27%	58,76%	11,99%	51,97%	76,86%	
7	Kosten	74.271,78 €	79.176,53 €	84.868,25 €	98.454,73 €	103.563,88 €	109.187,12 €	128.306,46 €	147.631,60 €	148.250,50 €	167.661,53 €	
8	Aufwand	174.880,64 €	187.919,81 €	210.808,69 €	317.259,85 €	353.265,60 €	375.869,14 €	379.596,84 €	206.646,55 €	405.064,65 €	597.231,21 €	
9	combined ratio	70,64%	71,20%	74,52%	96,67%	102,33%	103,27%	88,76%	41,99%	81,97%	106,86%	83,82%
10	Policenzahl	1.000	1.067	1.141	1.317	1.387	1.459	1.714	1.975	1.985	2.251	
11	Jahr	1	2	3	4	5	6	7	8	9	10	
34	23	0,00 €	0,00 €	0,00 €	0,00 €	0,00 €	3.706,44 €	0,00 €	0,00 €	0,00 €	0,00 €	





P / Z / Anz / Simulation / Q-Q-Plot

4. Literatur

4. Literatur

Bølviken, E. (2014): Computation and Modelling in Insurance and Finance. Cambridge University Press, Cambridge.

Kolonko, M. (2008): Stochastische Simulation. Grundlagen, Algorithmen und Anwendungen. Vieweg+Teubner, Wiesbaden.

Pfeifer, D. (2022): Monte Carlo Methoden.
http://www.staff.uni-oldenburg.de/dietmar.pfeifer/Monographien/Monte_Carlo_Methoden.pdf