Übungen zur Statistik: Fragen zu Verteilungen

♡ 1 **Beantworte die folgenden Statistik–Fragen zu Verteilungen auf der Grundlage der Skripte:**

(Verwende das Statistik-Skript und die Ergänzungen, zuerst wichtige fett gedruckt.)

(a) **Bernoulli-Verteilung:**
   i. Was ist das für eine Verteilung?
   ii. Eigenschaften?
   iii. Beispiele?

(b) **Binomial-Verteilung:**
   i. Was ist das für eine Verteilung?
   ii. Eigenschaften?
   iii. Beispiele?

(c) **Poisson-Verteilung:**
   i. Was ist das für eine Verteilung?
   ii. Eigenschaften?
   iii. Beispiele?

(d) Pascal-Verteilung:
   i. Was ist das für eine Verteilung?
   ii. Eigenschaften?
   iii. Beispiele?

(e) Geometrische Verteilung:
   i. Was ist das für eine Verteilung?
   ii. Eigenschaften?
   iii. Beispiele?

(f) **Hypergeometrische-Verteilung:**
   i. Was ist das für eine Verteilung?
   ii. Eigenschaften?
   iii. Beispiele?

(g) Rechtecksverteilung:
   i. Was ist das für eine Verteilung?
   ii. Eigenschaften?
   iii. Beispiele?

(h) **Normalverteilung:**
   i. Was ist das für eine Verteilung?

ii. Eigenschaften?
iii. Beispiele?

(i) *Erkläre folgende Sachverhalte und Verteilungen (mit Hilfe der Literatur und der Skripte):*

i. *Zum Grenzwertsätze von Moivre Laplace*
ii. *Lokaler Grenzwertsatz*
iii. *Grenzwertsatz von De Moivre/ Laplace*
iv. *Das Gesetz von Bernoulli der grossen Zahlen*
v. *Bemerkung zum Zufall*
vi. *Tschebyscheffsche Ungleichung*
vii. *Logarithmische Normalverteilung*
viii. *Exponentialverteilung*
ix. *Weibullverteilung*
x. *Gammaverteilung*
xi. *Ausblick*

♡ 2 Einige Erläuterungen:

(a) **Bernoulli-Verteilung:**

i. Um was handelt es sich?

Die Bernoulli–Verteilung ist eine „atomare“ Binomialverteilung mit $n = 1$ und $k = 0$ oder $k = 1$:

$$P(k) := P(X = k) := Be(k, p) = Bi(1, k, p), \;\; k \in \mathbb{N}_0, \;\; k \leq 1$$

ii. Siehe auch http://de.wikipedia.org/wiki/Bernoulli-Verteilung
Veranschaulichung: Galton-Brett, siehe
http://de.wikipedia.org/wiki/Galtonbrett

(b) **Binomial-Verteilung:**

i. Um was handelt es sich?

Sei z.B. $p = \frac{g}{m}$. Oder p sei als Grenzwert einer relativen Häufigkeit gegeben, $q = 1 - p$.

$$P(k) := P(X = k) := Bi(n, k, p) = \binom{n}{k} p^k \, q^{(n-k)}, \;\; n \in \mathbb{N}, \;\; k \in \mathbb{N}_0, \;\; k \leq n$$

In einem unendlich grossen Vorrat (Kiste, Urne) sind zwei Typen von Elementen A und B vorhanden. A kann immer mit der Wahrscheinlichkeit p und B mit der Wahrscheinlichkeit q gezogen werden.
Statt aus einem unendlichen Vorrat kann man auch aus einem endlichen Vorrat mit Zurücklegen ziehen.
Zieht man nun n Elemente, wovon k Elemente vom Typ A und der Rest, d.h. $n - k$ Elemente, vom Typ B sind, so multiplizieren sich der Unabhängigkeit wegen die Wahrscheinlichkeiten: $p \cdot p \cdot \ldots \cdot p \cdot q \cdot \ldots \cdot q = p^k \, q^{(n-k)}$. Dabei kann Typ A in der Reihenfolge auf $\binom{n}{k}$ Arten aus den n Elementen gezogen werden. Der Rest ist jeweils vom Typ B.

Es geht also hier um die Frage, was die Wahrscheinlichkeit ist bei n Zügen k mal Typ A zu ziehen und als Rest Typ B zu erhalten, wenn die Wahrscheinlichkeit für A und B bei jedem Zug dieselbe bleibt.

ii. Siehe auch http://de.wikipedia.org/wiki/Binomialverteilung
Veranschaulichung: Galton-Brett, siehe
http://de.wikipedia.org/wiki/Galtonbrett

(c) **Poisson-Verteilung:**

i. Um was handelt es sich?

Wird ein Bernoulli-Experiment sehr oft durchgeführt und ist die Erfolgswahrscheinlichkeit p sehr klein, so ist die Poisson–Verteilung eine gute Näherung für die entsprechende Binomial–Verteilung. Man bezeichnet daher die Poisson–Verteilung manchmal als die **Verteilung der seltenen Ereignisse**. Man sagt: Zufallsvariablen mit einer Poisson-Verteilung „genügen dem Poisson-Prozess". Druch eine Grenzwertbetrachtung oder eine differentielle Betrachtung gewinnt man mit $\lambda =$ Erwartungswert = Varianz:

$$P(k) := P(X = k) := P_\lambda(X = k) := Po(\lambda, k) = \frac{\lambda^k}{k!}\, e^{-k} = \lim_{n\to\infty} Bi(n, k, \frac{\lambda}{n}),\ \ p = \frac{\lambda}{n}$$

Beispiele: Die Poissonverteilung $P_\lambda(n)$ mit $\lambda = \dfrac{t_2}{t_1} = n \cdot p$ gibt die Wahrscheinlichkeit an, dass im Zeitraum t_2 genau n unabhängige Ereignisse stattfinden. λ ist „die mittlere Auftretenshäufigkeit eines Ereignisses".
Praktische Beispiele: Blitzeinschlagstatistik, Eintrittstatistik in ein Warenhaus, Statistik des radioaktiven Zerfalls u.s.w.

ii. Siehe auch http://de.wikipedia.org/wiki/Poisson-Verteilung

(d) **Pascal-Verteilung:**

i. Um was handelt es sich?

Die Pascal–Verteilung heisst auch **negative Binomialverteilung**.
Wie bei der Binomialverteilung wird aus einer Urne, in dem sich Objekte vom Typ A und vom Typ B befinden, mit Zurücklegen gezogen.
Dabei wird so lange ein Element von Typ A gezogen und wieder zurückgelegt, bis zum ersten Mal genau k solche Elemente vorhanden sind. Damit definieren wir eine Zufallsvariable $X :=$ „Zahl der Versuche, bis erstmals k Erfolge resultiert haben". k ist dabei vorgegeben. Daher variiert man die Zahl n der Versuche. X kann somit die Werte k, $k+1, \ldots, n_1, \ldots$ annehmen. X hat also abzählbar unendlich viele Ausprägungsmöglichkeiten.
Angenommen, bei $n-1$ Versuchen habe man $k-1$ Erfolge gehabt. Dann ist mit der Zufallsvariablen $Y :=$ „Zahl der Elemente Sorte A bei $n-1$ Versuchen":

$$P(Y = k-1) := Bi(n-1, k-1, p) = \binom{n-1}{k-1} p^{(k-1)}\, q^{(n-1-(k-1))}$$

Dann ist die Wahrscheinlichkeit, dass nochmals ein Element der Sorte A gezogen wird:

$$P(X = n) := P(Y = k-1) \cdot p = \binom{n-1}{k-1} p^k\, q^{(n-k)}$$

ii. Siehe auch http://de.wikipedia.org/wiki/Pascal-Verteilung .
(Beachte den Hinweis, wieso die Verteilung „negativ"genannt wird.)

(e) **Geometrische Verteilung:**

i. Um was handelt es sich?

Bei der geometrischen Verteilung geht es beim zugehörigen Bernoulli–Experiment um die Wahrscheinlichkeit, dass man bei n Ziehungen exakt beim $n+1$–ten Versuch Erfolg hat. Vorher hatte man n mal Misserfolg (eine Möglichkeit!):

$$P(Y=n) = p^1\, q^n = p\, q^n$$

Eine andere Definitionsart ist, dass man n Versuche bis zum ersten Erfolg benötigt, total also n Versuche:

$$P(X=n) = p^1\, q^{(n-1)} = p\, q^{(n-1)}$$

Anwendungen: Wartezeitenanalyse, Wahrscheinlichkeiten der Lebensdauer von Bauteilen u.s.w.
Die geometrische Verteilung ist ein Spezialfall der Pascal–Verteilung.

ii. Siehe auch http://de.wikipedia.org/wiki/Geometrische_Verteilung

(f) **Hypergeometrische-Verteilung:**

i. Um was handelt es sich?

Bei der hypergeometrischen Verteilung gehen wir wieder wie beim Bernoulli–Experiment von einer zweitypigen Grundgesamtheit (Typen A, B) aus. Wir ziehen zufällig n Elemente nacheinander ohne Zurücklegen.
Die hypergeometrische Verteilung gibt dann Auskunft darüber, mit welcher Wahrscheinlichkeit in der Stichprobe n_1 Elementen vom Typ A vorkommen. Gegeben sei eine Grundgesamtheit des Umfangs $N = n_1 + n_2$, n_1 Elemente sind vom Typ A, n_2 Elemente vom Typ B. M Elemente werden herausgegriffen. Wie gross ist die Wahrscheinlichkeit, dass man genau k Treffer (Typ A) herausgreift?

$$P(X=k) := Hyp(n_1, n_2, M, k) = \frac{g}{m} = \frac{\binom{n_1}{k} \cdot \binom{n_2}{M-k}}{\binom{n_1+n_2}{M}} = \frac{\binom{n_1}{k} \cdot \binom{n_2}{M-k}}{\binom{N}{M}}$$

Diese Verteilung ist in der Qualitätskontrolle bedeutend.

ii. http://de.wikipedia.org/wiki/Hypergeometrische_Verteilung

(g) **Rechteckverteilung oder stetige Gleichverteilung:**

i. Um was handelt es sich?

$$f(x) \quad = \quad \begin{cases} 0 & x \leq a \\ \dfrac{1}{b-a} & a < x < b \\ 0 & b \leq x \end{cases}$$

Bedeutung: Eine stetige Zufallsvariable X heisst gleichverteilt auf dem Intervall $[a, b]$, wenn eine Rechteckverteilung für $f(x)$. vorliegt.

Wichtig: Zufallszahlengeneratoren!

ii. Siehe auch http://de.wikipedia.org/wiki/Rechteckverteilung

(h) **Normalverteilung oder Gauss–Verteilung:**

i. Um was handelt es sich?

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

Dabei ist μ der Erwartungswert und σ die Standardabweichung.

Die besondere Bedeutung der Normalverteilung ist unter anderem durch den zentralen Grenzwertsatz bedingt. Dieser zeigt, dass eine Summe von n unabhängigen und identisch verteilten (d.h. ihre Verteilungen sind gleich) Zufallsvariablen im Grenzwert normalverteilt ist. Wenn also eine Zufallsvariable durch Überlagerung einer großen Zahl von Einflüssen entstehen, wobei jeder einzelne Einfluss einen im Vergleich zur Gesamtsumme unbedeutenden Beitrag liefert, so kann man sie als normalverteilt ansehen. Bei vielen Vorgängen des Ingenieurbereichs, aber auch der Natur– oder der Wirtschaftswissenschaften ist das in guter Näherung der Fall.

ii. Siehe auch http://de.wikipedia.org/wiki/Normalverteilung