Die Poissonverteilung

13

In diesem Kapitel

- Das Wahrscheinlichkeitsmodell von Poisson kennenlernen
- »Kleiner als«-, »größer als«- und andere Wahrscheinlichkeiten für die Poissonverteilung berechnen
- Den Erwartungswert und die Varianz der Poissonverteilung und den Poissonprozess berechnen
- Die Poissonverteilung mit einer Normalverteilung annähern

Mit der Poissonverteilung können Sie Fragen über die Zahl der Ankünfte oder Ereignisse beantworten, die innerhalb einer festgelegten Zeitspanne eintreten – etwa die Zahl der Flugzeuge, die innerhalb von zwei Stunden auf einem Flughafen landen oder die Zahl der Unfälle an einer Kreuzung innerhalb eines Monats. Mit der Poissonverteilung werden oft Geburts- und Todesraten modelliert, weil diese Raten Ereignisse repräsentieren, die innerhalb einer gegebenen Zeitspanne eintreten. Mit ihr können auch Vorkommnisse eines Ereignisses modelliert werden, die in einem festgelegten Raum auftreten, etwa die Zahl der Webfehler in einem 2 × 3-Meter-Teppich.

In diesem Kapitel lernen Sie, wie Sie Situationen erkennen, die sich mit der Poissonverteilung modellieren lassen, und wie Sie die entsprechenden Wahrscheinlichkeiten sowie den Erwartungswert und die Varianz der Poissonverteilung berechnen können. Außerdem erfahren Sie, wie Sie die Poissonverteilung mit einer Normalverteilung (siehe Kapitel 9) annähern können, wenn die Zahlen zu groß werden (dies ähnelt der Annäherung der Binomialverteilung an die Normalverteilung; siehe Kapitel 10).

Die Verteilung ist nach dem französischen Mathematiker und Physiker Siméon Denis Poisson (1781–1840) benannt, der dieses Modell erfand. Der Name wird etwa wie »Pwas-ong« ausgesprochen, wobei die Betonung auf dem »ong« liegt.

Ankünfte mit der Poissonverteilung modellieren

Mit der *Poissonverteilung* können Sie die Zahl der Ereignisse modellieren, die innerhalb einer festgelegten Zeitspanne oder einer angegebenen Fläche eintreten, und die zugehörigen Wahrscheinlichkeiten berechnen, etwa die Wahrscheinlichkeit, dass innerhalb von 15 Minuten mehr als zwei Telefonanrufe bei einer Helpline eingehen, dass Sie einen Schokoladenkeks ohne Schokoladenstückchen bekommen (was Sie bestimmt nicht wollen, oder?) oder dass eine Sekretärin mehr als zwei Tippfehler auf einer Seite macht und mehr.

Die Bedingungen für eine Poissonverteilung

Je mehr Wahrscheinlichkeitsmodelle Sie kennen, desto wichtiger wird es, auch die Bedingungen zu kennen, unter denen die einzelnen Modelle angewendet werden können. Wenn Sie ein Wahrscheinlichkeitsproblem mit einer Poissonverteilung lösen wollen, müssen Sie natürlich zunächst prüfen, ob dieses Modell für das Problem geeignet ist.

Eine Zufallsvariable X hat eine Poissonverteilung, wenn die folgenden Bedingungen erfüllt sind:

- ✔ X zählt die Ereignisse oder Vorkommnisse innerhalb einer festgelegten Zeit oder eines festgelegten Raumes.
- ✔ Die Ereignisse treten unabhängig voneinander ein.
- ✔ Zwei Ereignisse können nicht genau gleichzeitig eintreten.

Weil Sie die Ereignisse oder Vorkommnisse innerhalb einer festgelegten Zeit oder eines festgelegten Raumes zählen, kann die Zufallsvariable der Poissonverteilung jede positive Ganzzahl von null bis unendlich (0, 1, 2, 3 und so weiter) annehmen; deshalb ist eine Poisson-Zufallsvariable eine diskrete Zufallsvariable mit einer abzählbar unendlichen Zahl möglicher Werte (siehe Kapitel 7).

»Jetzt warten Sie mal«, sagen Sie. »Es ist nicht wirklich möglich, auf einer Seite eine unendliche Zahl von Tippfehlern zu machen oder innerhalb von 15 Minuten unendlich viele Telefonanrufe entgegenzunehmen.« Guter Einwand. Doch da es unmöglich ist, den genauen Cut-off-Punkt für solche Situationen anzugeben, lässt man X jede ganze Zahl von null bis unendlich annehmen. Aufgrund der Poisson-Wahrscheinlichkeitsmassenfunktion (WMF) wissen Sie, dass die Wahrscheinlichkeiten für extrem hohe Zahlen von Vorkommnissen klein sein werden (siehe den Abschnitt »Die Wahrscheinlichkeiten für die Poissonverteilung berechnen«).

Die Poisson- und die Binomialverteilung im Vergleich

Um das benötigte Wahrscheinlichkeitsmodell korrekt identifizieren zu können, müssen Sie den Unterschied zwischen den Wahrscheinlichkeitsmodellen der Poissonverteilung und der Binomialverteilung kennen (Näheres über die Binomialverteilung siehe Kapitel 8). Beide Verteilungen sind diskret, was bedeutet, sie können zum Zählen von Ergebnissen verwendet werden. Das Binomialmodell zählt die »Erfolge« (Ergebnisse mit einer gewünschten Eigenschaft) in n festgelegten Versuchen, wodurch X nur ganzzahlige Werte von null bis $n =$ die Zahl der Versuche annehmen kann. Dagegen hat die Poissonverteilung keine »Versuche«; sie beobachtet einfach eine Situation über eine festgelegte Zeitspanne oder in einem festgelegten Raum und zählt darin die interessierenden Vorkommnisse. Weil die Poissonverteilung keinen festgelegten Cut-off-Punkt hat, kann X alle ganzen Zahlen von null bis unendlich annehmen.

Ein Problem kann Szenarios für die Binomialverteilung und Poissonverteilung beschreiben, die sehr ähnlich klingen, weshalb Sie deren Unterschiede genau kennen müssen. Ein Beispiel: Sie sitzen an einer Kreuzung und beobachten den Verkehr. Sie sollen 50 Autos zufällig auswählen und zählen, wie viele vor dem Stoppschild der Kreuzung tatsächlich zum Stehen kommen. Welches Modell sollten Sie anwenden, die Binomialverteilung oder die Poissonverteilung? Weil die Zahl der Versuche festgelegt ist (50) und (bei einer zufälligen Stichprobe) jeder Versuch die gleiche Erfolgs- oder Misserfolgwahrscheinlichkeit hat, gilt hier das Binomialmodell (siehe Kapitel 8). Der Schlüssel zur Auswahl dieser Verteilung ist die festgelegte Zahl von Versuchen.

Ein weiteres Beispiel: Sie sitzen wieder an einer Kreuzung und beobachten den Verkehr. Jetzt sollen Sie zählen, wie viele Autos innerhalb von zwei Stunden das Stoppschild überfahren. Hier ist die Zeitspanne, aber nicht die Zahl der Versuche festgelegt. (Sie können gar nicht wissen, wie viele Autos an diese Kreuzung kommen werden.) In diesem Fall sollten Sie die Poissonverteilung verwenden.

Die Wahrscheinlichkeiten für die Poissonverteilung berechnen

Wenn Sie wissen, dass X einer Poissonverteilung folgt, können Sie mit den einschlägigen Formeln Wahrscheinlichkeiten für X berechnen. Für alle Poissonverteilungen gelten dieselben Formeln; der einzige Unterschied besteht in der Rate der Vorkommnisse, die Sie in einer Situation erwarten.

Es gibt zwei Möglichkeiten, Wahrscheinlichkeiten für eine Poissonverteilung zu berechnen: die Wahrscheinlichkeitsmassenfunktion und die kumulative Verteilungsfunktion.

Die WMF der Poissonverteilung

Die *Wahrscheinlichkeitsmassenfunktion* oder *WMF*, bezeichnet mit $P(x)$, (siehe Kapitel 7) liefert Ihnen die Formel für die Berechnung der Wahrscheinlichkeit, dass X gleich einer gewissen Zahl ist. Die Formel der WMF der Poissonverteilung lautet:

$$P(x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad \text{für} \quad x = 0,\ 1,\ 2,\ 3, \ldots,\ \infty$$

λ ist die sogenannte *Ereignisrate*, die durchschnittliche (oder mittlere) Rate des Eintretens des Ereignisses innerhalb einer festgelegten Zeitspanne oder eines festgelegten Raumes. (Das Problem muss Ihnen diesen Mittelwert geben.) Wenn Fachleute auf Daten stoßen, die zu diesem Modell passen, sagen sie, X habe eine Poissonverteilung mit der Ereignisrate λ.

Der Buchstabe e bedeutet in der Mathematik die Inverse des natürlichen Logarithmus (Inverse von ln). e^1 ist etwa 2,72. Die meisten wissenschaftlichen Taschenrechner verfügen über eine Taste für e^x; Sie können damit jede beliebige Potenz von e berechnen.

Ein Beispiel: Sie arbeiten an einem Helpdesk. Aus Erfahrung wissen Sie, dass Sie normalerweise durchschnittlich zehn Anfragen pro Stunde erhalten. Unter der Annahme, dass nicht zwei Anfragen zu genau demselben Zeitpunkt eingehen, hätte das Wahrscheinlichkeitsmodell für X – die tatsächliche Zahl der Anfragen pro Stunde – eine Poissonverteilung mit einer Ereignisrate λ = zehn pro Stunde. Die zugehörige WMF von X ist

$$P(x) = \frac{e^{-10}10^x}{x!}, \quad \text{für} \quad x = 0,\ 1,\ 2,\ 3, \ldots,\ \infty$$

Abbildung 13.1 zeigt den Graphen der WMF der Poissonverteilung mit der Ereignisrate λ.

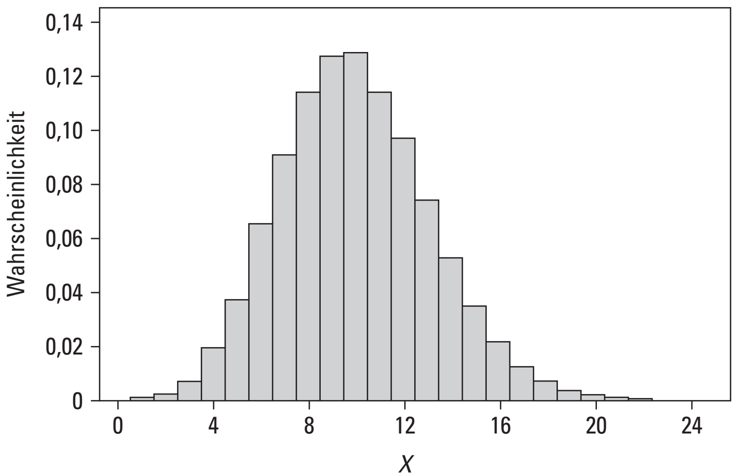

Abbildung 13.1: Eine Poissonverteilung mit der Ereignisrate $\lambda = 10$

Der Graph des WMF aus Abbildung 13.1 ist folgendermaßen zu interpretieren: Die Wahrscheinlichkeiten starten niedrig bei $X = 0$. Denn wenn Sie zehn Anfragen pro Stunde erwarten, ist die Wahrscheinlichkeit gering, nur wenige Anfragen zu erhalten. Dann nehmen die Wahrscheinlichkeiten mit X zu. Der Graph erreicht etwa bei zehn (die erwartete Zahl der Anfragen pro Stunde) einen Gipfel; dann werden die Werte immer kleiner, je größer X wird. Jenseits des Werts $X = 22$ sind fast keine Wahrscheinlichkeiten mehr vorhanden. Auch dies liegt daran, dass Sie zehn Anfragen pro Stunde erwarten; denn dann sollten mehr als 22 Anfragen in einer Stunde sehr selten vorkommen.

Sie können mit der WMF auch die Wahrscheinlichkeit berechnen, eine gewisse Anzahl von Anfragen pro Stunde zu bekommen. Nehmen Sie an, Sie suchten die Wahrscheinlichkeit, 15 Anfragen in der nächsten Stunde zu bekommen: $P(X = 15)$ oder $P(15)$. Wenn Sie 15 für X in die Formel für die WMF einsetzen, erhalten Sie

$$P(15) = \frac{e^{-10}10^{15}}{15!} = 0,035\ .$$

Wenn Sie die Wahrscheinlichkeit für 10 Anfragen pro Stunde suchen, setzen Sie 10 für X ein und erhalten

$$P(10) = \frac{e^{-10}10^{10}}{10!} = 0,125\ .$$

Wie Sie vielleicht erwarten, sollte die Wahrscheinlichkeit für Zahlen, die signifikant weit von 10 entfernt sind, klein sein. Denn wenn durchschnittlich 10 Anfragen pro Stunde erwartet werden, sollten sehr wenige oder sehr viele Anfragen nicht mit einer hohen Wahrscheinlichkeit eintreten. Zwei Beispiele: Die Wahrscheinlichkeit für 20 Anfragen pro Stunde beträgt

$$P(20) = \frac{e^{-10}10^{20}}{20!} = 0,002$$

und die Wahrscheinlichkeit für null Anfragen pro Stunde beträgt

$$P(0) = \frac{e^{-10}10^{0}}{0!} = 0,00005.$$

Vielleicht fragen Sie sich, warum die Wahrscheinlichkeit für $X = 10$ nicht höher als 0,125 ist, weil durchschnittlich 10 Anfragen erwartet werden. Obwohl dieser Wert von X die höchste Wahrscheinlichkeit hat (siehe Abbildung 13.1), variiert das Ergebnis von Stunde zu Stunde. Weil die Summe aller Wahrscheinlichkeiten eins sein muss (siehe Kapitel 7), verteilt sich die Wahrscheinlichkeit über verschiedene Werte von X, wobei der Wert der Ereignisrate λ die höchste Wahrscheinlichkeit hat. Je weiter Sie sich auf beiden Seiten von diesem Wert entfernen, desto geringer wird die Wahrscheinlichkeit. Die Poissonverteilung muss bei null beginnen, weshalb sie nicht symmetrisch ist; deshalb sieht die Wahrscheinlichkeitsverteilung auch schief aus, weil sie in einigen Situationen sozusagen keine Zeit hat, auf der linken Seite einen Schwanz auszubilden.

Bei Poisson-Wahrscheinlichkeitsmodellen müssen Sie auf die Einheiten achten. In dem vorhergehenden Beispiel gilt die Ereignisrate $\lambda = 10$ für die erwartete Zahl der Anfragen pro Stunde, und Sie müssen dieses »pro Stunde« in Ihrer Beschreibung von X berücksichtigen. (In dem Abschnitt »Zeitliche oder räumliche Einheiten ändern: der Poissonprozess« weiter hinten in diesem Kapitel wird beschrieben, wie Wahrscheinlichkeiten für eine Poissonverteilung berechnet werden, wenn die Einheiten geändert wurden.)

Die KVF der Poissonverteilung

Die *kumulative Verteilungsfunktion* oder *KVF*, bezeichnet mit F(x), (siehe Kapitel 7) liefert Ihnen die Formel für die Berechnung der kumulierten Wahrscheinlichkeit von null bis zu einem speziellen Wert von X. Mit der KVF können Sie die Wahrscheinlichkeit berechnen, dass X kleiner oder größer als ein Wert ist oder zwischen zwei Werten liegt. Die Formel der KVF der Poissonverteilung lautet

$$\sum_{X \leq x} \frac{e^{-\lambda} \lambda^X}{X!}, \quad \text{für} \quad x = 0,\ 1,\ 2,\ 3, \ldots,\ \infty$$

Die Formel summiert die Werte der WMF (siehe den vorherigen Abschnitt) von null bis zu einem gewünschten Wert x. Beispielsweise zeigt Abbildung 13.2 die KVF einer Poissonverteilung mit der Ereignisrate $\lambda = 10$. Sie beginnt für $X = 0$ bei null und nähert sich dem Wert eins, wenn X gegen unendlich geht.

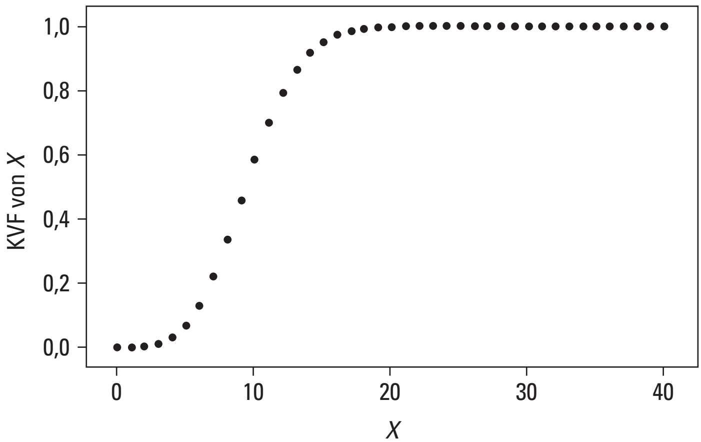

Abbildung 13.2: Die KVF einer Poissonverteilung mit der Ereignisrate $\lambda = 10$

Sie müssen Werte von $F(x)$ nicht mit der Formel für die KVF berechnen, sondern können die Werte in einer Tabelle nachschlagen (siehe Tabelle A.3 im Anhang).

Um Werte von $F(x)$ für eine Poissonverteilung mit der Poissontabelle zu ermitteln, müssen Sie nur zwei Dinge kennen:

- ✔ Den Wert von λ
- ✔ Den Wert von x, bis zu dem die Wahrscheinlichkeiten kumuliert werden sollen

Die gesuchte Wahrscheinlichkeit befindet sich im Schnittpunkt der Zeile mit dem X-Wert und der Spalte mit dem λ-Wert.

»Kleiner als oder gleich«- oder »kleiner als«-Wahrscheinlichkeiten berechnen

Wenn Sie eine »kleiner als oder gleich«-Wahrscheinlichkeit suchen, müssen Sie alle Wahrscheinlichkeiten für die Werte von X addieren, die kleiner als die oder gleich der gewünschten Zahl sind. Ein Beispiel: Sie suchen die Wahrscheinlichkeit, dass X kleiner als oder gleich vier ist, also: $P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$. Dies ist dasselbe, wie den Wert der KVF, geschrieben $F(4)$, zu suchen. Die Tabelle erspart Ihnen die Additionsarbeit. Sie enthält »kleiner als oder gleich«-Wahrscheinlichkeiten für eine Poissonverteilung für verschiedene Werte von λ. In diesem Abschnitt lernen Sie, wie Sie diese Tabelle (Tabelle A.3 im Anhang) anwenden können.

Die Poissontabelle besteht aus zwei Teilen: Der erste Teil enthält Werte für λ kleiner als oder gleich eins, der zweite Werte für λ größer als eins (bis 20).

Ein Beispiel: Ein Teppichhersteller möchte die Qualität seiner Teppiche kontrollieren. Aus Erfahrung weiß er, dass er einen Mangel pro 100 Quadratmeter Teppich erwarten darf. In diesem Beispiel ist also $\lambda = 1$ Mangel pro 100 qm. Wie hoch ist die Wahrscheinlichkeit, höchstens zwei (zwei oder weniger) Mängel pro 100 qm zu finden: $P(X \leq 2)$ oder $F(2)$? Es liegt eine Poissonverteilung mit einer Ereignisrate $\lambda = 1$ vor. Suchen Sie im oberen Teil der Poissontabelle die Spalte für $\lambda = 1$ und die Zeile für $X = 2$. Im Schnittpunkt dieser Zeile und Spalte finden Sie die Wahrscheinlichkeit 0,920. Die Wahrscheinlichkeit, dass ein 100 qm großer Teppich zwei oder weniger Mängel aufweist, beträgt 0,920 oder 92 Prozent.

Mit der Poissontabelle können Sie strikte »kleiner als«-Wahrscheinlichkeiten ermitteln. Da diese Wahrscheinlichkeiten nicht in der Tabelle enthalten sind, müssen Sie das Problem umformulieren. Wenn Sie beispielsweise die Wahrscheinlichkeit suchen, dass X kleiner als vier ist, können Sie (da die X-Werte ganzzahlig sind) auch die Wahrscheinlichkeit suchen, dass X kleiner als oder gleich drei ist: $P(X < 4) = F(3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)$.

Ein Beispiel: Die Wahrscheinlichkeit, weniger als zwei Mängel in 100 qm Teppich zu finden, ist $P(X < 2)$. Da die Poissontabelle nur Wahrscheinlichkeiten für »kleiner als oder gleich« enthält, formulieren Sie das Problem um: $P(X < 2) = P(X \leq 1)$. In der Tabelle finden Sie dafür die Wahrscheinlichkeit 0,736. Dies ist die Wahrscheinlichkeit, weniger als zwei Mängel in 100 qm Teppich zu finden.

»Größer als«-Wahrscheinlichkeiten berechnen

Mit der Poissontabelle (siehe Tabelle A.3 im Anhang) können Sie auch »größer als«-Wahrscheinlichkeiten ermitteln. Da die Tabelle nur »kleiner als oder gleich«-Wahrscheinlichkeiten enthält, müssen Sie das Problem mit dem Komplement der gesuchten Wahrscheinlichkeit entsprechend umformulieren. Ein Beispiel: Die Wahrscheinlichkeit, dass X größer als sechs ist, ist dieselbe wie eins minus die Wahrscheinlichkeit, dass X kleiner als oder gleich sechs ist. Letztere Wahrscheinlichkeit finden Sie in der Poissontabelle. Denken Sie einfach daran: Wenn »größer als« gesucht ist, müssen Sie »eins minus« verwenden.

Um bei dem Teppichbeispiel aus dem vorhergehenden Abschnitt zu bleiben: Wie hoch ist die Wahrscheinlichkeit, mehr als zwei Mängel in 100 qm Teppich zu finden: $P(X > 2)$? Um die Poissontabelle nutzen zu können, bilden Sie das Komplement: $P(X > 2) = 1 - P(X \leq 2) = 1 - 0{,}920 = 0{,}08$. Die Wahrscheinlichkeit, mehr als zwei Mängel in 100 qm Teppich zu finden, beträgt also 8 Prozent.

Analog ist die Wahrscheinlichkeit, mindestens zwei Mängel in 100 qm Teppich zu finden, $P(X \geq 2) = 1 - P(X \leq 1) = 1 - 0{,}736 = 0{,}264$ oder 26,4 Prozent.

Wahrscheinlichkeiten zwischen zwei Werten ermitteln

Wenn Sie die Wahrscheinlichkeit suchen, dass X zwischen zwei Werten liegt, müssen Sie unterscheiden, ob die Grenzen des Intervalls ein- oder ausgeschlossen werden sollen. So bedeutet $P(3 < X < 6)$ die Wahrscheinlichkeit, dass alle Werte von X zwischen 3 (ausschließlich) und 6 (ausschließlich) liegen – anders ausgedrückt: die Wahrscheinlichkeit $P(X = 4 \text{ oder } X = 5)$. Andererseits bedeutet $P(3 \leq X \leq 6)$ die Wahrscheinlichkeit, dass alle Werte von X zwischen 3 (einschließlich) und 6 (einschließlich) liegen – anders ausgedrückt: die Wahrscheinlichkeit $P(X = 3 \text{ oder } X = 4 \text{ oder } X = 5 \text{ oder } X = 6)$. Natürlich können Sie die Relationen auch kombinieren: $P(3 \leq X < 6)$, was 3 (einschließlich) und 6 (ausschließlich) bedeutet, oder $P(3 < X \leq 6)$, was 3 (ausschließlich) und 6 (einschließlich) bedeutet.

Um die Wahrscheinlichkeit zu ermitteln, dass X zwischen zwei Werten liegt, müssen Sie das Problem so umformulieren, dass es nur »kleiner als oder gleich«-Wahrscheinlichkeiten enthält, dann die Werte der KVF für jede Zahl aus der Tabelle ablesen und die Wahrscheinlichkeit des kleineren Werts von der des größeren Werts subtrahieren.

Um bei dem Teppichbeispiel aus den beiden vorhergehenden Abschnitten zu bleiben: Wie hoch ist die Wahrscheinlichkeit, bei 100 qm Teppich zwischen drei und fünf Mängel zu finden: $P(3 \leq X \leq 5)$? Zunächst lesen Sie die Wahrscheinlichkeit $F(5)$ für alle Werte bis fünf aus der Poissontabelle ab. Dann subtrahieren Sie die Wahrscheinlichkeit $F(2)$ für alle Werte bis zwei. Die Differenz ist die Wahrscheinlichkeit aller Werte von X zwischen 3 und 5. Laut Tabelle sind $F(5) = 0{,}999$ und $F(2) = 0{,}920$. Sie erhalten $F(5) - F(2) = 0{,}999 - 0{,}920 = 0{,}079$ oder 7,9 Prozent. Mithin beträgt die Wahrscheinlichkeit, bei 100 qm Teppich zwischen drei und fünf Mängel zu finden, 7,9 Prozent.

Der Erwartungswert und die Varianz der Poissonverteilung

Der Erwartungswert der Poissonverteilung sagt Ihnen, wie viele Ereignisse Sie innerhalb einer festgelegten Zeitspanne oder einem festgelegten Raum erwarten können. Ihre Varianz sagt Ihnen, mit welcher Variabilität dieser Ergebnisse Sie von einem Experiment zum nächsten rechnen müssen. Das Interessante an der Poissonverteilung ist die sehr enge Beziehung zwischen dem Erwartungswert und der Varianz.

Der *Erwartungswert* einer Verteilung ist der Gesamtdurchschnittswert der Verteilung. Die Notation für den Erwartungswert ist μ_x oder $E(X)$ oder bei der Poissonverteilung $\mu_x = E(X) = \lambda$.

Die Problemstellung muss Ihnen die Ereignisrate λ der Poissonverteilung nennen. Möglicherweise wird sie nicht als *Ereignisrate* oder *Erwartungswert* bezeichnet, sodass Sie sich an diese Tatsache erinnern müssen. Möglicherweise sagt Ihnen das Problem auch nicht speziell, was λ ist, sondern spricht von einem Mittelwert »so und so«. Dieses »so und so« repräsentiert den Wert von λ.

Die *Varianz* einer Verteilung ist der quadrierte Gesamtdurchschnitt der Abstände vom Mittelwert. Sie wird mit σ_x^2 bezeichnet. Die *Standardabweichung* ist die Quadratwurzel aus der Varianz. Sie wird mit σ_x bezeichnet. Bei der Poissonverteilung ist die Varianz $\sigma_x^2 = \lambda$. Weil dadurch die Einheiten in Quadrat-Einheiten umgewandelt werden, ist die Bedeutung nicht immer einsichtig (etwa wenn es um die Zahl der Telefonanrufe pro Minute geht), deshalb wird die Variabilität der Ergebnisse normalerweise mit der Standardabweichung gemessen. Die Standardabweichung einer Poissonverteilung ist

$$\sigma_x = \sqrt{\lambda}.$$

Ein Beispiel: Die Kunden einer Bank treffen mit einer Ereignisrate $\lambda = 20$ pro Stunde ein; ihr Eintreffen folgt einer Poissonverteilung. In diesem Falle ist der Erwartungswert $\mu_x = \lambda = 20$ Kunden pro Stunde, und die Standardabweichung ist $\sigma_x = \sqrt{\lambda} = \sqrt{20} = 4{,}47$ Kunden pro Stunde.

Denken Sie immer an die Einheiten des Mittelwerts einer Poissonverteilung. Genau genommen ist der Mittelwert eine Rate pro Einheit, und diese Einheit ist wichtig und sollte immer angegeben werden – insbesondere wenn sich die Einheiten ändern. (Im nächsten Abschnitt finden Sie mehr über das Thema der Einheiten.)

Zeitliche oder räumliche Einheiten ändern: der Poissonprozess

Der *Poissonprozess* ist ein Wahrscheinlichkeitsmodell, das von einer Poissonverteilung ausgeht und deren zeitliche oder räumliche Einheiten sich ändern. Er verwendet eine Funktion, mit der Sie leicht die Einheiten einer Poissonverteilung ändern und sofort Wahrscheinlichkeiten, Erwartungswerte und Varianzen ermitteln können. Ein Beispiel: X sei eine Zufallsvariable mit einem Mittelwert von λ pro Einheit. Eine neue Zufallsvariable $Y = \alpha X$ sei ein Poissonprozess mit der Ereignisrate $\alpha\lambda$, wobei α die Zahl ist, mit der Sie λ multiplizieren müssen, um die alten Einheiten in die neuen umzurechnen. Anders ausgedrückt: Eine Änderung der Einheiten für X wird in den Formeln für den Erwartungswert und die Varianz berücksichtigt.

Ein Beispiel: X sei die Zahl der Unfälle pro Jahr; die Ereignisrate sei $\lambda = 1$ pro Jahr. Dann zählt $2X$ die Zahl der Unfälle in zwei Jahren; die Ereignisrate beträgt dann $2\lambda = 2 \cdot 1 = 2$ pro zwei Jahre. Deshalb gilt hier $\alpha = 2$.

Manchmal können Sie den benötigten Multiplikator leicht identifizieren, manchmal ist es schwierig. Aber Sie können immer ein Verhältnisproblem (eine Dreisatzaufgabe!) konstruieren, um den Multiplikator zu ermitteln. (Erinnern Sie sich an Ihren Matheunterricht? Und Sie glaubten, Sie könnten dies nie verwenden …) Ein Beispiel: X sei eine Poissonverteilung mit einer Ereignisrate von 16 Ereignissen pro Stunde; Sie wollen allerdings nur Intervalle von jeweils 15 Minuten betrachten. Welchen Multiplikator α benötigen Sie, um eine Rate pro Stunde in eine Rate pro 15 Minuten umzurechnen? Ein einfacher Dreisatz hilft Ihnen weiter: Sie wissen, dass eine Stunde aus 60 Minuten besteht. Deshalb können Sie fragen: Wenn 16 Ereignisse in 60 Minuten erwartet werden, wie viele Ereignisse α können Sie dann in 15 Minuten erwarten? Oder mathematisch formuliert: $16/60 \text{ min} = \alpha/15 \text{ min}$. Wenn Sie diese Gleichung nach α auflösen, erhalten Sie: $\alpha = 16 \cdot 15\,\text{min}/60\,\text{min} = 16/4 = 4$. Das Ergebnis ist plausibel: Wenn Sie durchschnittlich 16 Ereignisse pro Stunde erwarten, können Sie durchschnittlich mit etwa 4 Ereignissen in 15 Minuten (einer Viertelstunde) rechnen.

Sie können den Poissonprozess auf das weiter oben eingeführte Teppichbeispiel anwenden: Es wurden 200 qm Teppich produziert. Die Zahl Y der Mängel in 200 qm hat eine Poissonverteilung mit einer Ereignisrate $2\lambda = 2 \cdot 1 = 2$. Sie können diesen Wert nun in eine Poissonverteilung einsetzen. Der Erwartungswert von Y ist $\mu_y = 2$; die Varianz ist gleich ${\sigma_y}^2 = \lambda = 2$ Mängel pro 200 qm Teppich (zum Quadrat); und die Standardabweichung ist gleich $\sigma_y = \sqrt{\lambda} = \sqrt{2} = 1{,}41$ Mängel pro 200 qm Teppich. (Die Berechnung wird im vorhergehenden Abschnitt erklärt.)

Die Standardabweichung (1,41) von Y, die Zahl der Mängel pro 200 qm, ist nicht dasselbe wie das Zweifache der Standardabweichung (1) von X, der Zahl der Mängel pro 100 qm. Der Unterschied kommt dadurch zustande, dass die Mängel bei 200 qm mehr Raum haben, um einzutreten, wodurch die Variabilität, wann und wo sie eintreten, größer ist. Aus der Tatsache, dass pro 200 qm durchschnittlich zwei Mängel auftreten, dürfen Sie nicht schließen, dass diese Mängel fein säuberlich auf separate 100-qm-Stücke verteilt sind. Deshalb dürfen Sie nicht einfach die Ergebnisse der ursprünglichen Poissonverteilung verdoppeln, um die Antworten für einen Poissonprozess mit einem Multiplikator von zwei zu berechnen.

Nehmen Sie jetzt an, Sie untersuchten nur 50 qm Teppich. Wenn die ursprüngliche Zufallsvariable X die Zahl der Mängel pro 100 qm und Y die Zahl der Mängel pro 50 qm ist, erhalten Sie laut Poissonprozess $Y = 1/2 \cdot X$. Also ist der Multiplikator $\alpha = 1/2$ oder 0,50. Sie können sagen, dass Y eine Poissonverteilung mit einer Ereignisrate ist, die gleich der Hälfte der Ereignisrate der ursprünglichen Poissonverteilung X beträgt, sodass es $1/2 \cdot 1 = 1/2$ Mängel pro 50 qm gibt. Der Erwartungswert von Y ist ein halber Mangel pro 50 qm; die Varianz ist 1/2

Mangel pro 50 qm (zum Quadrat); und die Standardabweichung ist $\sqrt{\frac{1}{2}} = 0,71$ Mängel pro 50 qm. Die Wahrscheinlichkeit, höchstens zwei Mängel pro 50 qm zu finden, ist $P(X \leq 2) =$ 0,986 (laut Tabelle A.3 im Anhang im Schnittpunkt der Zeile für $X = 2$ und der Spalte für $\lambda = 0{,}50$).

Eine Poissonverteilung an eine Normalverteilung annähern

Bei der Berechnung von Wahrscheinlichkeiten für eine Poissonverteilung stoßen Sie oft auf eine Grenze, die die Ermittlung der Werte der WMF und / oder der KVF schwierig macht. Ein Beispiel: Gegeben sei eine Poissonverteilung mit einer Ereignisrate von 50. Sie suchen die Wahrscheinlichkeit, dass $X = 45$ ist:

$$P(X = 45) = \frac{e^{-50}50^{45}}{45!}$$

Falls Sie Teile dieser Aufgabe mit einem Taschenrechner lösen wollen, werden Sie kein Glück haben. Deshalb können Sie sich vorstellen, dass die Berechnung der KVF für 45, die die Summe aller Wahrscheinlichkeiten für X von 0 bis 45 ergeben würde, eine ebenso unmögliche Aufgabe wäre.

Weil die Berechnungen für eine Poissonverteilung für große Werte von λ extrem schwierig sind, müssen Sie oft mit einer Annäherung arbeiten. So wie Sie unhandliche Binomialverteilungen an die Normalverteilung annähern können (siehe Kapitel 10), können Sie Gleiches mit der Poissonverteilung tun. Sie müssen nur herausfinden, welche Werte Sie für μ, den Mittelwert der angenäherten Normalverteilung, und σ, ihre Standardabweichung, verwenden müssen. In diesem Abschnitt wird die Annäherung der Poissonverteilung an die Normalverteilung behandelt.

Die Bedingungen einer Annäherung an die Normalverteilung erfüllen

Eine Poissonverteilung muss gewisse Bedingungen erfüllen, damit eine Annäherung an die Normalverteilung möglich ist. Zu diesem Zweck muss die Poissonverteilung über eine Ereignisrate $\lambda = 20$ oder mehr verfügen. Je größer der Wert von λ, desto genauer sind die Wahrscheinlichkeiten, die die Annäherung der Poissonverteilung an die Normalverteilung liefert. Mit der Annäherung an die Normalverteilung ist es erheblich einfacher, Wahrscheinlichkeiten für die Poissonverteilung zu ermitteln, als mit der WMF-Formel.

Für Werte von λ über 20 enthält die Poissontabelle (siehe Tabelle A.3 im Anhang) keine Werte für die KVF (siehe den Abschnitt »Die Wahrscheinlichkeiten für die Poissonverteilung berechnen« weiter vorne in diesem Kapitel); die Experten mussten die Tabelle irgendwo begrenzen, und bei den meisten Poisson-KVF-Tabellen wird dafür der Wert $\lambda = 20$ gewählt.

Um zu verstehen, warum eine Annäherung einer Poissonverteilung an die Normalverteilung in einigen Situationen brauchbar ist, sollten Sie sich den Graphen der Wahrscheinlichkeitsmassenfunktion einer Poissonverteilung ansehen (mehr über die WMF finden Sie in dem Abschnitt »Die Wahrscheinlichkeiten für die Poissonverteilung berechnen« weiter vorne in diesem Kapitel).

Abbildung 13.3 zeigt ein Histogramm einer Poissonverteilung mit einer Ereignisrate von $\lambda = 2$. Die linke Seite des Graphen scheint abgeschnitten zu sein und auf der rechten Seite zeigt sich ein Schwanz. Der Graph wurde so gezeichnet, weil die Ereignisrate (der Mittelwert) zwei beträgt und deshalb sehr nahe bei null liegt. Auf der linken Seite des Graphen gibt es nicht genug Platz, um viele Wahrscheinlichkeiten zu zeigen (weil X bei einer Poissonverteilung nicht kleiner als null sein kann). Dadurch ist der Graph der WMF rechtsschief. (In Kapitel 4 steht Näheres über die Form von Verteilungen.) Die Annäherung an die Normalverteilung funktioniert hier nicht gut, weil der Wert von λ zu klein ist.

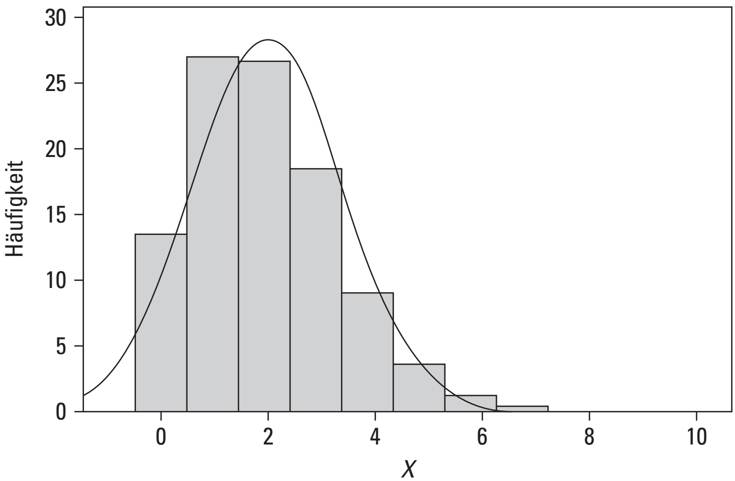

Abbildung 13.3: Eine Poissonverteilung, deren Ereignisrate $\lambda = 2$ für eine Annäherung an die Normalverteilung zu klein ist

Abbildung 13.4 zeigt eine Poissonverteilung mit einer Ereignisrate von $\lambda = 5$. Sie können bereits sehen, dass sich der Graph von null weg verschiebt und sich auf der linken Seite ein Schwanz ausbildet. Doch der Graph ist immer noch rechtsschief, weil der Mittelwert immer noch recht nahe bei null liegt, was bedeutet, dass die Annäherung an die Normalverteilung auch hier nicht gut funktioniert, weil der Wert von λ immer noch zu klein ist.

Abbildung 13.5 zeigt eine Poissonverteilung mit einer Ereignisrate von $\lambda = 20$. Jetzt können Sie sehen, wie sich eine glockenförmige Kurve herausbildet, weil die linke Seite vollständig zur Verfügung steht, da die Ereignisrate weit genug von null entfernt ist, sodass der gesamte linke Schwanz dargestellt werden kann. An diesem Punkt können Sie problemlos die Annäherung der Poissonverteilung an die Normalverteilung verwenden.

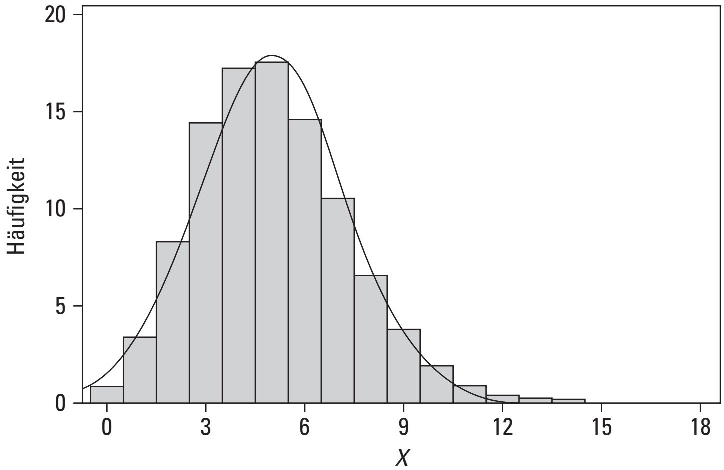

Abbildung 13.4: Eine Poissonverteilung, deren Ereignisrate $\lambda = 5$ für eine Annäherung an die Normalverteilung immer noch zu klein ist

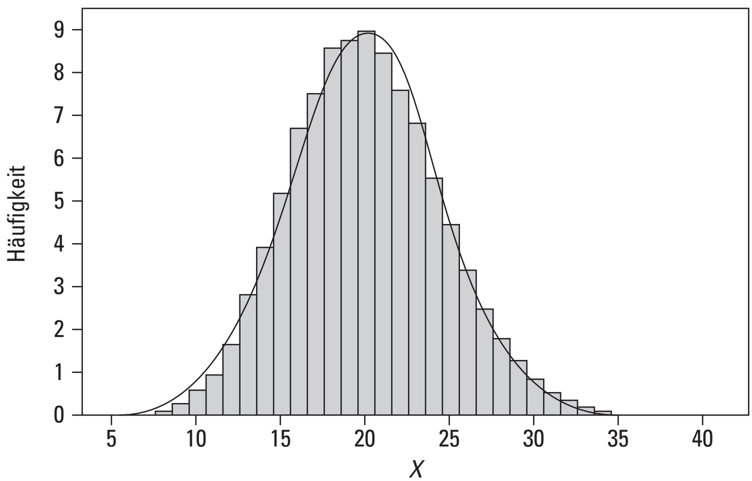

Abbildung 13.5: Eine Poissonverteilung, deren Ereignisrate $\lambda = 20$ für eine Annäherung an die Normalverteilung hinreichend groß ist

Die vollständigen Schritte für die Annäherung der Poissonverteilung an die Normalverteilung

Wenn die Ereignisrate λ einer Poissonverteilung größer als 20 ist, und Sie Wahrscheinlichkeiten für die Poissonverteilung mit der Annäherung der Poissonverteilung an die Normalverteilung berechnen wollen, führen Sie die Schritte aus Kapitel 10 aus, um eine Wahrscheinlichkeit mit einer Normalverteilung annähernd zu berechnen. Zu diesem Zweck müssen Sie den Mittelwert und die Standardabweichung der Normalverteilung kennen.

Hier verwenden Sie einfach die Ereignisrate und die Standardabweichung der Poissonverteilung: $\mu_x = \lambda$ beziehungsweise $\sigma_x = \sqrt{\lambda}$.

Die Wahrscheinlichkeiten einer Poissonverteilung werden wie folgt mit einer Normalverteilung angenähert:

1. **Prüfen Sie, ob der Wert von λ hinreichend groß ist (je größer, desto besser, aber mindestens 20).**
2. **Übersetzen Sie das Problem in die Wahrscheinlichkeitsnotation: $P(X < a)$, $P(X > b)$ oder $P(a < X < b)$.**
3. **Transformieren Sie a und / oder b mit der Z-Formel in einen Z-Wert: $Z = (X - \mu) / \sigma$.**

   Bei der Poissonverteilung verwenden Sie $\mu_x = \lambda$ für den Mittelwert und $\sigma_x = \sqrt{\lambda}$ für die Standardabweichung.
4. **Lesen Sie die zugehörigen Wahrscheinlichkeiten aus der Z-Tabelle ab (Tabelle A.2 im Anhang).**
5. **Wenn Sie eine »kleiner als oder gleich«-Wahrscheinlichkeit suchen, sind Sie fertig. Wenn Sie eine »größer als oder gleich«-Wahrscheinlichkeit suchen, berechnen Sie das Komplement (eins minus die Wahrscheinlichkeit aus der Z-Tabelle, Schritt 4). Wenn Sie eine Wahrscheinlichkeit zwischen zwei Werten suchen, führen Sie die Schritte 2 bis 4 für beide Werte aus und subtrahieren die Wahrscheinlichkeiten (die größere minus die kleinere).**
6. **Beantworten Sie die ursprüngliche Frage im Kontext des Problems (das heißt in der Sprache von X, nicht der von Z). Die Antwort bleibt dieselbe.**

Ein Beispiel: Die Kunden einer Bank betreten den Schalterraum mit einer durchschnittlichen Rate von fünf Kunden pro zehn Minuten. Wie hoch ist die Wahrscheinlichkeit, dass mehr als 35 Kunden innerhalb einer Stunde den Schalterraum betreten (achten Sie auf die Änderung der Einheiten)? Dieses Beispiel ist ein Poissonprozess mit dem Multiplikator sechs (denn sechsmal zehn Minuten ergibt eine Stunde, die neue Zeiteinheit; siehe den Abschnitt »Zeitliche oder räumliche Einheiten ändern: der Poissonprozess«). Die Zahl der Kunden in einer Stunde sei X. X hat eine Poissonverteilung mit der Ereignisrate $\lambda = 6 \cdot 5 = 30$. (Sie wissen auch, dass die Standardabweichung $\sqrt{\lambda} = \sqrt{30} = 5{,}48$ Kunden pro Stunde beträgt.) Sie suchen die Wahrscheinlichkeit, dass X größer als 35 ist: $P(X > 35)$.

Mit den folgenden Schritten ermitteln Sie diese Wahrscheinlichkeit durch eine Annäherung an eine Normalverteilung:

1. **$\lambda = 30$ ist für eine Annäherung an eine Normalverteilung hinreichend groß.**
2. **Sie suchen $P(X > 35)$.**
3. **Sie setzen den Mittelwert $\mu_x = \lambda = 30$ und die Standardabweichung $\sigma_x = \sqrt{\lambda}$ (die Ereignisrate beziehungsweise die Standardabweichung der Poissonverteilung) in die Z-Formel ein und erhalten $Z = \frac{X - \mu_x}{\sigma_x} = \frac{35 - 30}{\sqrt{30}} = \frac{5}{5{,}48} = 0{,}91$.**

   Also entspricht $P(X > 35)$ etwa $P(Z > 0{,}91)$.

4. **Sie lesen die Wahrscheinlichkeit für 0,91 aus der Z-Tabelle (Tabelle A.2 im Anhang) ab und erhalten 0,8186.**

   Weil die Z-Tabelle die Wahrscheinlichkeit aller Werte kleiner als der Z-Wert liefert, wissen Sie, dass $P(Z < 0{,}91) = 0{,}8186$ ist.

5. **Da Sie eine »größer als«-Wahrscheinlichkeit suchen, müssen Sie $P(Z > 0{,}91)$ berechnen. Mit dem Komplement erhalten Sie: $1 - P(Z < 0{,}91) = 1 - 0{,}8186 = 0{,}1814$.**

6. **Im ursprünglichen Kontext des Problems formulieren Sie Ihre Antwort: Die Wahrscheinlichkeit, dass mehr als 35 Kunden die Schalterhalle innerhalb einer Stunde betreten, beträgt etwa 0,1814 oder 18,14 Prozent.**

Vergessen Sie nicht, dass Ihre endgültige Antwort ein Näherungswert ist, wenn Sie mit einer Annäherung der Poissonverteilung an die Normalverteilung arbeiten. Dies sollte in der Formulierung Ihrer Antwort auch zum Ausdruck kommen. Sie können auch sagen: »Ich vertraue dieser Annäherung, weil λ größer als 20 ist«.

Wenn die Werte von λ über 20 liegen, liefert die Normalverteilung ziemlich genaue Annäherungen an die wahren Wahrscheinlichkeiten der Poissonverteilung, aber bis etwa $\lambda = 40$ gibt es immer noch einen merklichen Unterschied. Sie können dieses Problem mit einer Stetigkeitskorrektur beheben, damit Ihre Antworten mit der Normalverteilung noch näher an den genauen Antworten liegen, die Sie mit der Poissonverteilung erhalten würden. Zu diesem Zweck müssen Sie nur 0,50 von Ihrem X-Wert subtrahieren, wenn Sie eine »größer als oder gleich«-Wahrscheinlichkeit suchen beziehungsweise 0,50 zu Ihrem X-Wert addieren, wenn Sie eine »kleiner als oder gleich«-Wahrscheinlichkeit suchen, bevor Sie die Z-Formel anwenden. (Stetigkeitskorrekturen werden ausführlich in Kapitel 10 behandelt.) In dem Beispiel aus diesem Abschnitt suchen Sie $P(X \geq 36)$. Mit Stetigkeitskorrektur suchen Sie mit der Z-Formel tatsächlich nach $P(X \geq 35{,}5)$. Sie erhalten dann 0,1587 statt 0,1814.

Die Poissonverteilung in Aktion

Es gibt zahlreiche praktische Anwendungen für die Poissonverteilung. Beispielsweise werden Warteschlangen mit diesen Verteilungen simuliert. Lebensversicherungen berechnen damit Geburts- und Todesraten, um ihre Beiträge festzulegen sowie zu ermitteln, wann die Leistungen ausgezahlt werden müssen. Straßenverkehrsbehörden modellieren damit das Unfallaufkommen verschiedener Verkehrswege, um deren Gefährlichkeit abzuschätzen.

Zu langweilig? Nun, mit der Poissonverteilung wird auch gewährleistet, dass Ihr morgendlicher Fruchtjoghurt Stückchen enthält, die zumindest wie Früchteteile aussehen. Der Hersteller setzt die Ereignisrate der Poissonverteilung, mit der die Zahl der farbigen Stückchen in dem Joghurt gemessen wird, so hoch an, dass die Wahrscheinlichkeit, dass keine Stückchen in dem Joghurt sind, fast null ist. Was für ein Trost!