Serie 03 - Solutions

3.1 Anwendungsbeispiel 1 Wahrscheinlichkeitsdichtefunktion

> **Frage 1**: Sind die Werte im Intervall $[0, 60]$ Anzahl CHF?
>
> **Frage 2**: Was repräsentiert die y-Achse?

> **Notiz 1**: Die x-Achse repräsentiert einen Werte, wie zum Beispiel Menge oder CHF.
>
> **Notiz 2**: Die Wahrscheinlichkeit P, dass meine Werte im Intervall $[a, b]$ liegen, kann wie folgt dargestellt werden: $P(a < X < b) = \int_a^b f(x)dx = Wahrscheinlichkeit$
>
> **Notiz 3**: Es ist wichtig, dass die gesammte Fläche unter der Funktion 1 ergibt.

a.) Konstanter Wert c berechnen

$$f(x) = \begin{cases} cx(15 - \frac{x}{4}) & \text{falls } 0 \leq x \leq 60 \\ 0 & \text{sonst} \end{cases}$$

> Die Konstante c ist ein unbekannter Wert, der uns verhindert x zu berechnen. Deshalb müssen wir ihn aus der folgenden Bedingung berechnen:
>
> - Wir wissen, dass die gesammte Fläche unterhalb der Funktion 1 geben muss, denn eine Wahrscheinlichkeit kann nicht über 1 (100%) sein.
> - Wir wissen, dass die Kurve im Intervall $[0, 60]$ liegt, denn überall sonst ist die Wahrscheinlichkeit 0

Wir rechnen das Integral der Wahrscheinlichkeitsdichtefunktion $f(x) = cx(15 - \frac{x}{4})$ im Intervall $[0, 60]$ aus und setzen dieses gleich 1:

$$F(x) = \int_0^{60} f(x)\, dx = \int_0^{60} cx(15 - \frac{x}{4})\, dx = c \cdot \int_0^{60} 15x - \frac{x^2}{4}\, dx = 1$$

$$F(x) = c \cdot \int_0^{60} 15x - \frac{x^2}{4}\, dx = c \cdot \left| \frac{15}{2}x^2 - \frac{x^3}{12} \right|_0^{60} = c \cdot \left(\frac{15}{2}(60)^2 - \frac{(60)^3}{12} \right) - 0 = 1$$

$$F(x) = c \cdot \left(\frac{15}{2}(60)^2 - \frac{(60)^3}{12} \right) - 0 = c \cdot 9000 - 0 = 1$$

$$c = \frac{1}{9000}$$

Nun haben wir die Konstante c berechnet und können weitere Berechnungen mit der nun vollständigen Wahrscheinlichkeitsdichtefunktion durchführen.

b.) Verteilungsfunktion $F(X)$ der Zufallsvariablen X angeben

Da wir nun c berechnet haben können wir die Verteilungsfunktion problemlos angeben. Die Zufallsvariablen X werden als Parameter übergeben, da wir die Wahrscheinlichkeit für das Auftreten der Zufallsvariable X im Intervall $[a, b]$ berechnen wollen. Da wir das Integral von $f(x)$ im Interval $[a, b]$ berechnet haben müssen wir nur die Funktion mit den verschiedenen Wertebereichen angeben.

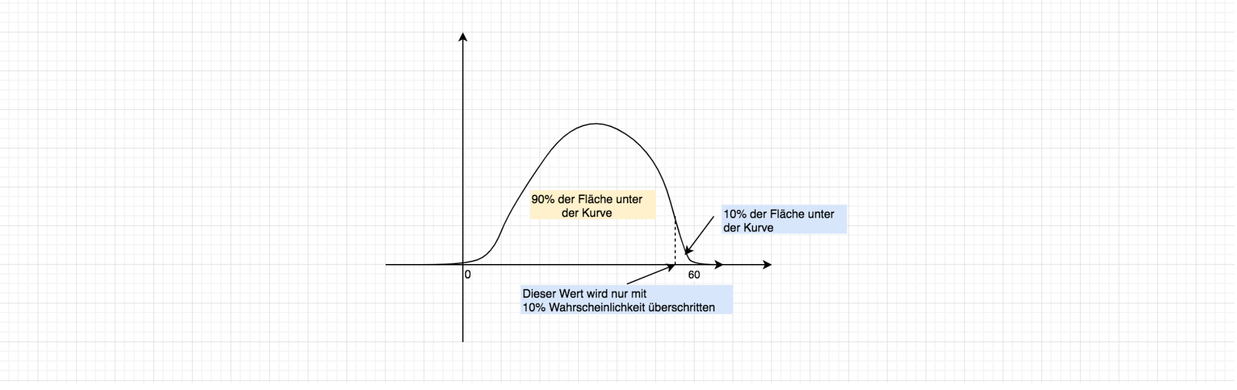
$$F(x) = \frac{1}{9000} \left(\frac{15}{2}x^2 - \frac{x^3}{12} \right) \quad \text{falls } 0 \leq x \leq 60$$

Und wenn wir nun alle Wertebereiche in Betracht ziehen (auch diese ausserhalb $0 \leq x \leq 60$), dann erhalten wir:

$$F(x) = \begin{cases} 0 & \text{falls } 0 < x \\ F(x) = \frac{1}{9000} \left(\frac{15}{2}x^2 - \frac{x^3}{12} \right) & \text{falls } 0 \leq x \leq 60 \\ 1 & \text{falls } x > 60 \end{cases}$$

c.) Welcher Wert der Aufwendungen mit 10% Wahrscheinlichkeit überschritten?

Wir können nun die Verteilungsfunktion $F(x)$ verwenden, um den Wert der Aufwendung mit dieser Wahrscheinlichkeit zu berechnen.



Dies bedeutet, dass wir einen Wert für x (in diesem Fall ist x die monatliche Aufwendung für den Wasserverbrauch eines 2-Personenhaushalts in CHF) finden müssen, welcher 90% der gesammten Wahrscheinlichkeit (Fläche unter der Kurve) entspricht. Mathematisch gesehen bedeutet das (wenn a den gesuchte Wert darstellt):

$$F(a) = \frac{1}{9000}\left(\frac{15}{2}a^2 - \frac{a^3}{12}\right) = 0.9$$

Jetzt löst man das mit dem Taschenrechner (oder mit WolframAlpha) auf und erhält für a folgende Werte:

- $a = -28.625$
- $a = 48.252$
- $a = 70.373$

> Man erhält 3 Werte, weil die Verteilungsfunktion ein Polynom dritten Grades ist.

Wir nehmen nun den einzigen Wert der sich in unserem Interval $[0, 60]$ befindet und zwar $a = 48.252$.

d.) Erwarteten monatlichen Aufwendungen berechnen

> **Frage 3**: Wieso wird hier das Intervall $[0, 60]$ verwendet in der Lösung und nicht $[-\infty, \infty]$ wie in der Formel?
>
> **Frage 4**: Wann weiss man, dass es sich bei einer Wahrscheinlichkeitsdichtefunktion um eine normalverteilte symmetrische Kurve handelt?

In dieser Aufgabe liegt die Antwort, wie wir vorgehen müssen, bereits in der Aufgabenstellung. Die **erwarteten** monatlichen Aufwendungen sind zu berechnen. Dabei handelt es sich, um den wahrscheinlichsten Wert oder dem Schwerpunkt der Funktion: **der Erwartungswert**.

Wir kennen bereits die Formel, um den Erwartungswert einer **stetigen** Wahrscheinlichkeitskurve zu berechnen:

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\, dx$$

Wie man sieht wird hier nicht die Verteilungsfunktion $F(x)$ verwendet, sondern die Wahrscheinlichkeitsdichtefunktion $f(x)$. In unserem Beispiel heisst das folgendes:

$$E(X) = \int_0^{60} x \cdot cx \left(15 - \frac{x}{4}\right) = \int_0^{60} x \cdot \frac{1}{9000} x \left(15 - \frac{x}{4}\right) = \frac{1}{9000} \int_0^{60} x^2 \left(15 - \frac{x}{4}\right)$$

$$E(X) = \frac{1}{9000} \cdot \left| 5x^3 - \frac{1}{16} x^4 \right|_0^{60} = \frac{1}{9000} \cdot (1080000 - 810000) = 30$$

> **Alternative**: Da es sich bei der Wahrscheinlichkeitsdichtefunktion um eine normalverteilte Gaussche Glockenkurve handelt, weiss man, dass der Schwerpunkt direkt in der Mitte der Kurve liegt und die Dichte um diesen Wert symmetrisch verteilt ist. Somit kann man auch die Hälfte von 60 nehmen und den Erwartungswert direkt berechnen.

3.2 Anwendungsbeispiel 2 Wahrscheinlichkeitsdichtefunktion

a.) Warum handelt es sich bei c um den Wert 0.1 ?

> Wir wissen, dass
>
> - die Fläche unter der Kurve immer gleich 1 sein muss, da nicht über eine 100% -ige Wahrscheinlichkeit erziehlen kann. Die Kurven sind auf 100% normiert.
> - die Formel, um die Flächen von Dreiecken zu berechnen: $A = \frac{a \cdot b}{2}$

Mit unserem Vorwissen können wir nun folgendes berechnen:

$$A = \frac{c \cdot 20}{2} = 1$$

Der Wert c ist somit gleich $\frac{1}{10}$ oder auch als 0.1 geschrieben. Nun müssen wir noch die Dichtefunktion $f(x)$ finden. Dies machen wir folgendermassen.

- $\text{Steigung der Kurve } = -\frac{0.1}{20}$
- $\text{Geradenfunktion f(x) } = y = mx + q$

Damit erhalten wir die Wahrscheinlichkeitsdichtefunktion $f(x) = \frac{1}{10}(1 - \frac{x}{20})$ für den Bereich $[0, 20]$. Für alle Wertebereiche lautet die Dichtefunktion:

$$f(x) = \begin{cases} 0 & \text{falls } 60 < x < 20 \\ \frac{1}{10}(1 - \frac{x}{20}) & \text{falls } 20 \le x \le 60 \end{cases}$$

b.) Wahrscheinlichkeit, dass Wert kleiner als 5 und 10 ist, berechnen

Wir haben hier die beiden Werte 5 und 10 gegeben, die Werte auf der x-Achse beschreiben. Da wir die Dichtefunktion in der vorherigen Teilaufgabe ermittelt haben, können wir diese hier verwenden, um die kummulative Verteilungsfunktion $F(x)$ von X zu berechnen. Diese lässt sich berechnen durch Integration der Dichtefunktion $f(x)$.

$$F(x) = \int_0^x f(t)\, dt = \int_0^x \frac{1}{10}\left(1 - \frac{t}{20}\right)\, dt = \frac{1}{10}\left(x - \frac{x^2}{40}\right) = \frac{x}{10} - \frac{x^2}{400}$$

> Merke: Dies gilt nur für den Wertebereich $[0, 20]$. Für die Werte wo $x \leq 0$ ist, ist die Wahrscheinlichkeit 0 und für die Werte wo $x \geq 20$ ist, ist die Wahrscheinlichkeit 1.

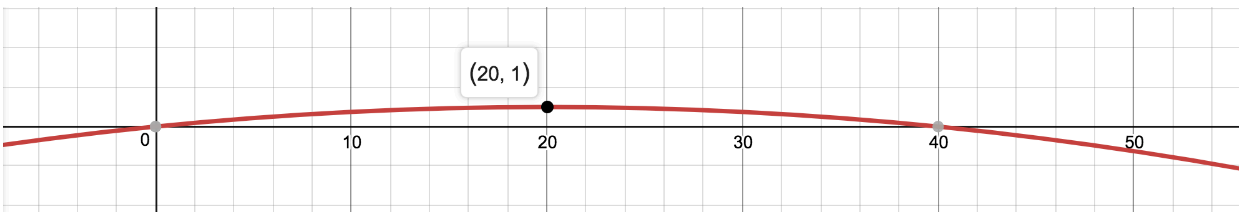
- $F(5) = \frac{(5)}{10} - \frac{(5)^2}{400} = 0.4375 = 43.75\%$
- $F(10) = \frac{(10)}{10} - \frac{(10)^2}{400} = 0.75 = 75\%$

Die Wahrscheinlichkeit, dass Werte unter 5 angenommen werden, liegt bei 43.75% und dass Werte unter 10 angenommen werden liegt bei 75%. Wir schreiben:

- $P[X < 5] = 43.75\%$
- $P[X < 10] = 75\%$

c.) Kummulative Verteilungsfunktion skizzieren

Am besten skizziert man sich die gesammte Funktion (ohne Rücksicht auf das Intervall $[0, 20]$) und schneidet dann eifach bei 20 ab. So sieht man, dass die Funktion bis 20 steigt.



d.) Erwartungswert, Median und Standardabweichung berechnen

Erwartungswert

Den Erwartungswert können wir wie im Beispiel vorhin mit der Wahrscheinlichkeits-dichtefunktion $f(x)$ berechnen. Die Formel dazu lautet:

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\, dx$$

Nun setzen wir die zuvor ermittelte Wahrscheinlichkeitsdichtefunktion in diese Formel ein, um den Erwartungswert zu berechnen:

$$E(X) = \int_0^{20} x \cdot \left(\frac{1}{10}(1 - \frac{x}{20})\right)\, dx = \int_0^{20} x \cdot \left(\frac{1}{10} - \frac{x}{200}\right)\, dx = \left|\frac{x^2}{20} - \frac{x^3}{600}\right|_0^{20}$$

$$E(X) = \left|\frac{x^2}{20} - \frac{x^3}{600}\right|_0^{20} = \left[\frac{(20)^2}{20} - \frac{(20)^3}{600}\right] - 0 = \frac{20}{3} = 6.\overline{666}$$

Median

Der Median entspricht 50% der Fläche unter der Wahrscheinlichkeitsdichtefunktion. Deshalb wissen wir, dass die kummulative Verteilungsfunktion $F(x) \stackrel{!}{=} 0.5$ sein muss im Bereich $[0, 20]$.

$$F(x) = \frac{x}{10} - \frac{x^2}{400} \stackrel{!}{=} 0.5$$

Wir wandeln die Funktion um damit wir ein Polynom zweiten Grades erhalten und diese dann mit der "Mitternachtsformel" oder eine andere Methode eurer Wahl (z.B. WolframAlpha) auflösen können.

$$\frac{x^2}{400} - \frac{x}{10} + 0.5 \stackrel{!}{=} 0$$

Nach Auflösen dieser Gleichung erhalten wir die folgenden zwei Resultate:

- $x = 5.85786$
- $x = 34.1421$

Da der zweite Wert nicht in unserem gesuchten Wertebereich liegt können wir diesen vernachlässigen und erhalten als Median den Wert 5.85786.

Standardabweichung

> Wir wissen, dass

> - Die Standardabweichung ist gleich der Wurzel der Varianz $\sigma_x = \sqrt{Var(X)}$
> - Die Varianz lässt sich aus dem Erwartungswert berechnen: $Var(X) = E(X^2) - (E(X))^2$

Aus dem vorhin berechneten Erwartungswert können wir nun die Varianz berechnen, um damit nachher die Standardabweichung ausrechnen zu können. Zuerst berechnen wir jedoch die Teilkomponenten der Varianzformel:

$$E(X)^2 = (\frac{20}{3})^2 = (6.\overline{666})^2 = \frac{400}{9}$$

$$E(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x^2)\, dx = \int_0^{20} x^2 \cdot \frac{1}{10}(1 - \frac{x^2}{20}) = \frac{200}{3}$$

$$Var(X) = E(X^2) - (E(X))^2 = \frac{200}{3} - \frac{400}{9} = \frac{200}{9}$$

Die Standardabweichung lässt sich nun kann einfach aus der Wurzel der Varianz ($\frac{200}{9}$) berechnen:

$$\text{Standardabweichung } \sigma_x = \sqrt{Var(X)} = \sqrt{\frac{200}{9}} = 4.7140$$

e.) Wahrscheinlichkeit, dass Kosten höchstens $120'000.-\text{Fr.}$ sind

Wir haben die Formel um aus der Zufallsvariable X die Kosten K zu berechen. Diese lautet $K = 40000 \cdot \sqrt{X}$. Wir wollen die Wahrscheinlichkeit kennen, bei der die Kosten $K = 120000.-\text{Fr.}$ sind. Um diese Wahrscheinlichkeit ausrechnen zu können, müssen wir zuerst die Zufallsvariable X kennen.

$$K = 120000 = 40000 \cdot \sqrt{X}$$

$$\text{Zufallsvariable } X = 9$$

> Wir wissen, dass
>
> - wir für die Wahrscheinlichkeit die kummulative Verteilungsfunktion brauchen
> - die kummulative Verteilungsfunktion $F(x) = \frac{x}{10} - \frac{x^2}{400}$ (siehe Teilaufgabe b.))

> **Merke**: Wir wollen nur die Fläche bis und mit der Zufallsvariable kennen, denn in der Aufgabenstellung steht "höchstens" 120'000 Fr. und nicht "mindestens".

Nun können wir diese Zufallsvariable X in die kummulative Verteilungsfunktion einsetzen und erhalten die entsprechende Wahrscheinlichkeit.

$\displaystyle F(9) = P[X \leq 9] = \frac{(9)}{10} - \frac{(9)^2}{400} = 0.6975 = 69.75\%$

f.) Für welchen Parameter λ hat die Exponentialverteilung denselben Erwartungswert?

> **Frage 5**: Wieso kann man einfach eine andere Verteilung nehmen? Verändert dies nicht die komplette Aussage der Wahrscheinlichkeiten?
>
> Wir wissen, dass
>
> - Der Erwartungswert in einer Exponentialverteilung wie folgt definiert ist: $E(X) = \frac{1}{\lambda}$
> - Der Erwartungswert dieser Aufgabe $\frac{20}{3}$ ist.

Wir nehmen den zuvor berechneten Erwartungswert $E(X) = \frac{20}{3}$ und verwenden die Formel $E(X) = \frac{1}{\lambda}$ um den Parameter λ zu berechnen.

$E(X) = \frac{1}{\lambda} = \frac{20}{3}$

$\lambda = \frac{3}{20}$

g.) Wahrscheinlichkeit mit Exponentialverteilung berechnen

> Wir wissen, dass
>
> - $\lambda = \frac{3}{20}$ ist
> - die Formel für die kummulate Verteilungsfunktion mit einer Exponentialverteilung ist immer $F(x) = 1 - e^{-\lambda x}$, wobei sich der Parameter λ sich je nach Werten / Wahrscheinlichkeiten ändert
> - die Zufallsvariable $X = 9$

Wir berechnung nun mit der kummulativen Verteilungsfunktion einer Exponentialverteilung die Wahrscheinlichkeit erneut wie in der Teilaufgabe e.):

$F(9) = 1 - e^{-(\frac{3}{20})(9)} = 0.7408 = 74.08\%$

Nehmen wir an die Dauer der Baustellen ist exponentialverteilt, ist die Wahrscheinlichkeit, dass die Kosten einer Baustelle unter $120000.-\mathbf{Fr.}$ liegt grösser. Die verschiedenen Verteilungen sind einfach Modelle, um die Wahrscheinlichkeiten annehmen zu können. Aus Erfahrung nimmt man für bestimmte Anwendungsbeispielen unterschiedliche Verteilungen an. Merke jedoch wie die Erwartungswerte immer gleich bleiben.

> **Frage 6**: Stimmt diese Aussage oben?

3.3 Exponentialverteilung

> Wir wissen, dass
>
> - der Parameter $\lambda = 0.04$
> - der Median 50% der Fläche unter der Wahrscheinlichkeitsdichtefunktion
>
> - der Wertebereich über 0 ist $\rightarrow W_X = [0, \infty)$
> - die Wahrscheinlichkeitsdichtefunktion $f(x) = \lambda e^{-\lambda x}$
> - die kummulative Verteilungsfunktion $F(x) = 1 - e^{-\lambda x}$
> - der Erwartungswert $E(X) = \frac{1}{\lambda}$
> - die Varianz $Var(X) = \frac{1}{\lambda^2}$

a.) Median, Erwartungswert und Wahrscheinlichkeit berechnen

Median

Wir kennen λ (0.04) und wissen, dass der Median 50% der Fläche unter der Wahrscheinlichkeitsdichtefunktion entspricht und können somit folgendes gleichstellen:

$$F(x) \stackrel{!}{=} 1 - e^{-(0.04)x} \stackrel{!}{=} 0.5$$

$$x = 17.3$$

Hier entspricht x dem Median, denn bei diesem Wert befinden sich die Hälfte der Werte unterhalt und die andere Hälfte oberhalb dieses Wertes.

Erwartungswert

Der Erwartungswert ist ziemlich einfach zu ermitteln, denn wir kennen das folgenden Gesetzt $E(X) = \frac{1}{\lambda}$ für exponentialverteilte Modelle.

$$E(X) = \frac{1}{(0.04)} = 25h$$

Der Erwartungswert entspricht in dieser Aufgabe der Lebenserwartung einer Maschine. Wir können jetzt mit der kummulativen Verteilungsfunktion berechnen, mit welcher Wahrscheinlichkeit die Maschine mindestens die Lebenserwartung erreicht $\rightarrow 1 - P[T_1 > E(T_1)]$. Dies entspricht gleich der Wahrscheinlichkeit des Erwartungswertes.

$E(25) = ... \rightarrow TODO$

> **Frage 7**: Wieso entspricht dieser Wert nicht gleich dem Wert des Erwartungswertes?