



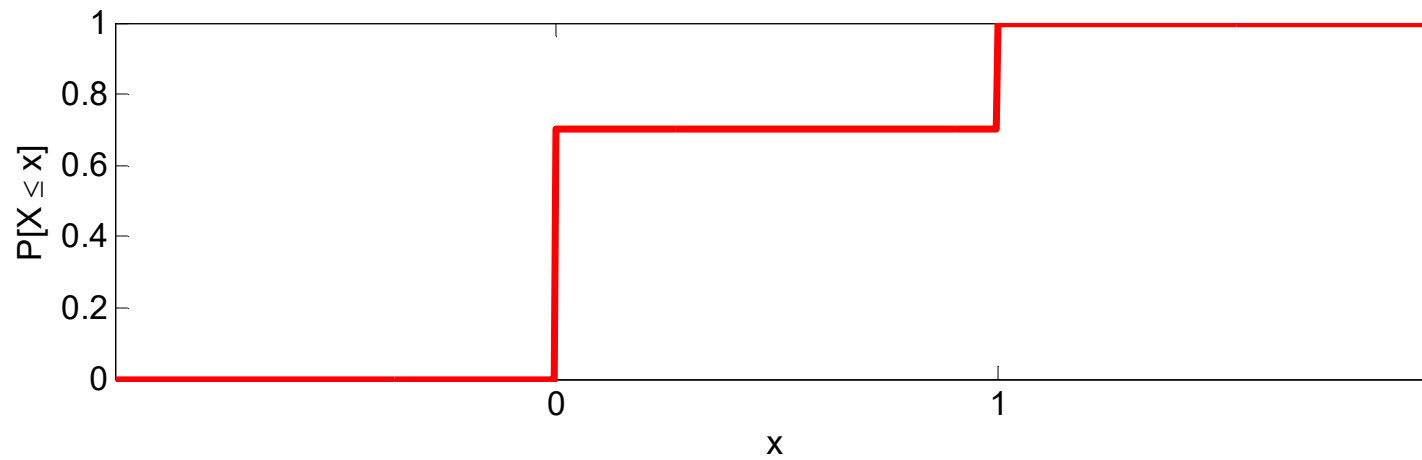
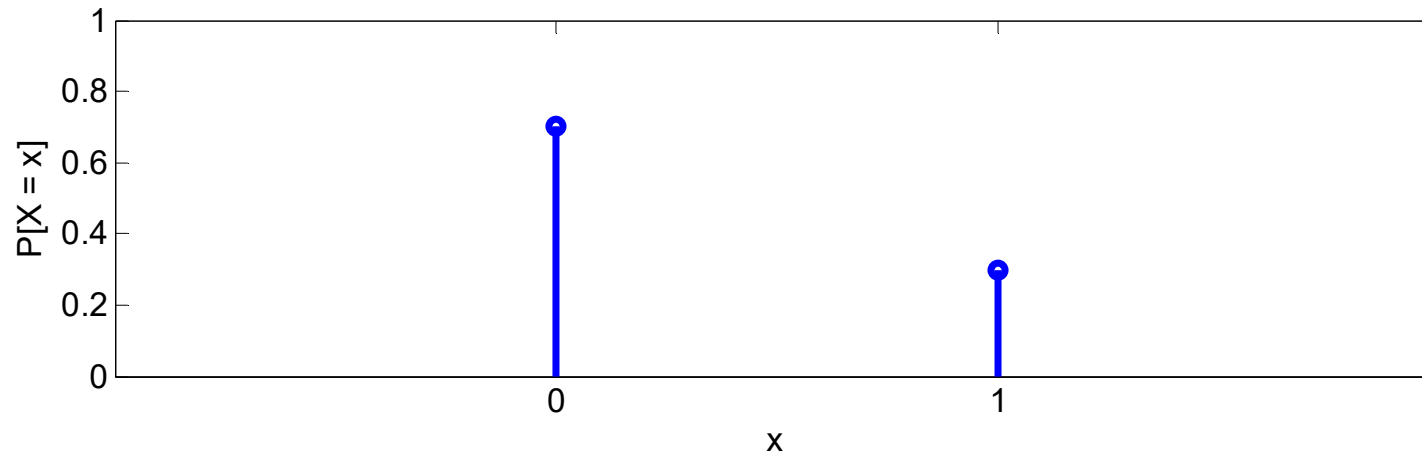
Diskrete Verteilungen

Bernoulli-Verteilung: $X \sim \text{Bernoulli}(s)$

Symbol für «verteilt wie»

- «Experiment» mit **zwei** Ausgängen: «Erfolg» ($X = 1$) oder «Misserfolg» ($X = 0$). Die Erfolgswahrscheinlichkeit sei s .
- Wertebereich: $W = \{0,1\}$
- Wa'keitsfunktion:
$$P(X = 0) = 1 - p; \quad P(X = 1) = p$$
- Erwartungswert / Varianz (nachrechnen!)
$$E(X) = p$$
$$\text{Var}(X) = p(1 - p)$$

Bernoulli-Verteilung ($p = 0.3$)



Bernoulli-Verteilung: Beispiele

«Experiment» und «Erfolg» können vieles bedeuten:

- Erdbeben tritt ein ($X = 1$) vs. Erdbeben tritt nicht ein ($X = 0$).
- Qualitätsanforderung erfüllt ($X = 1$) vs. Qualitätsanforderung nicht erfüllt ($X = 0$) (bzw. umgekehrt).
- Gesund ($X=1$) vs. krank ($X=0$).
- Usw...

Also immer wenn etwas **zwei mögliche Ausgänge** hat, kann man die Bernoulli-Verteilung verwenden.

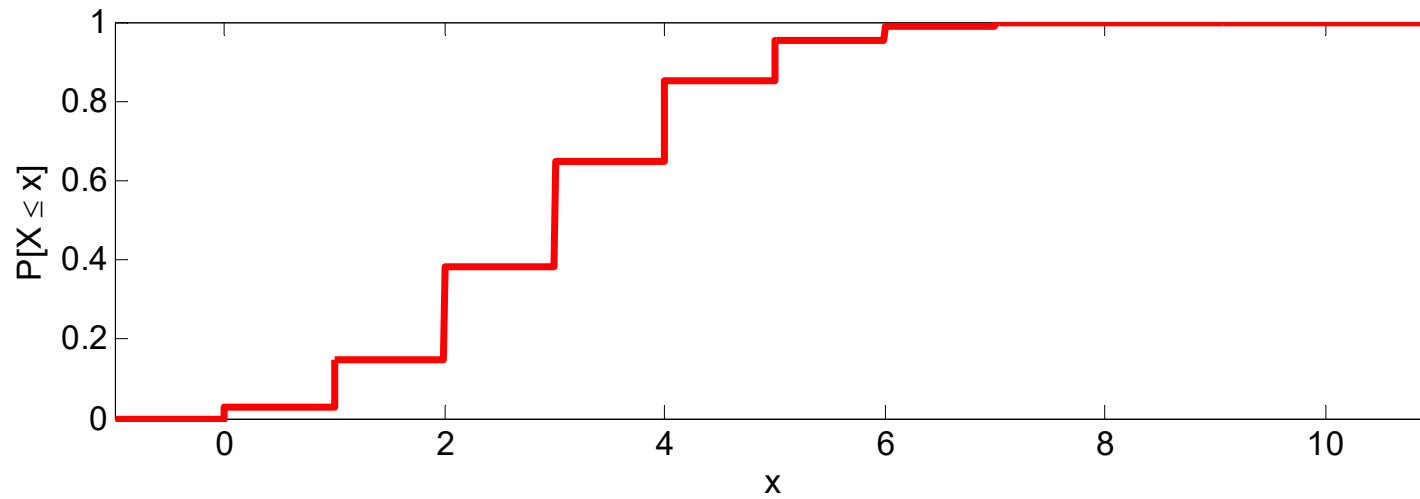
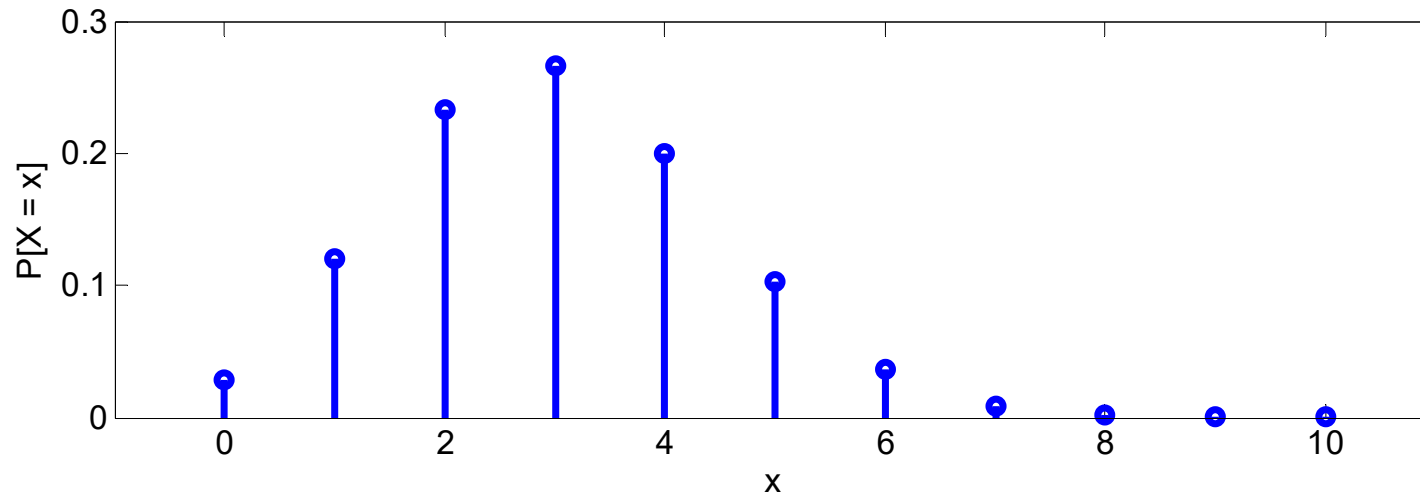
Binomialkoeffizient

- Wir werden jetzt n unabhängige Experimente betrachten und wir zählen die Anzahl Erfolge.
- Wieviel Möglichkeiten gibt es für k Erfolge in n Experimenten?
- Dies ist gegeben durch den **Binomialkoeffizient** $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.
- Beispiel: Anzahl Möglichkeiten für 1 Erfolg in 5 Experimenten:
10000, 01000, 00100, 00010, 00001 und
$$\binom{5}{1} = \frac{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{4 \cdot 3 \cdot 2 \cdot 1 \cdot 1} = 5$$
- Beispiel: Anzahl Möglichkeiten für 2 Erfolge in 4 Experimenten:
1100, 1010, 1001, 0110, 0101, 0011 und
$$\binom{4}{2} = \frac{4 \cdot 3 \cdot 2 \cdot 1}{2 \cdot 1 \cdot 2 \cdot 1} = \frac{4 \cdot 3}{2} = 6$$

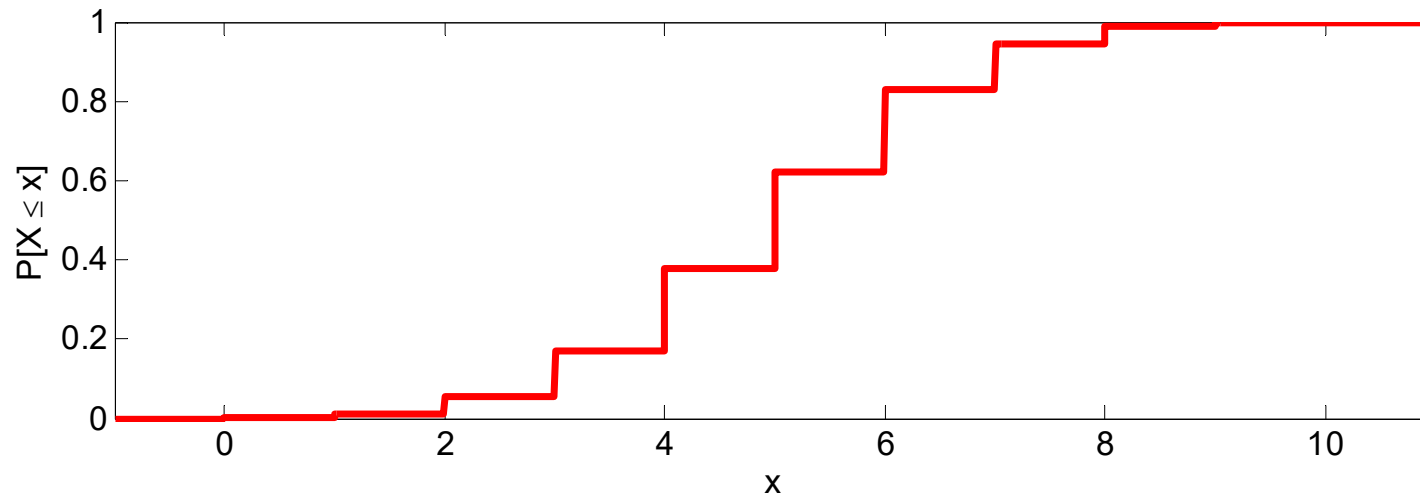
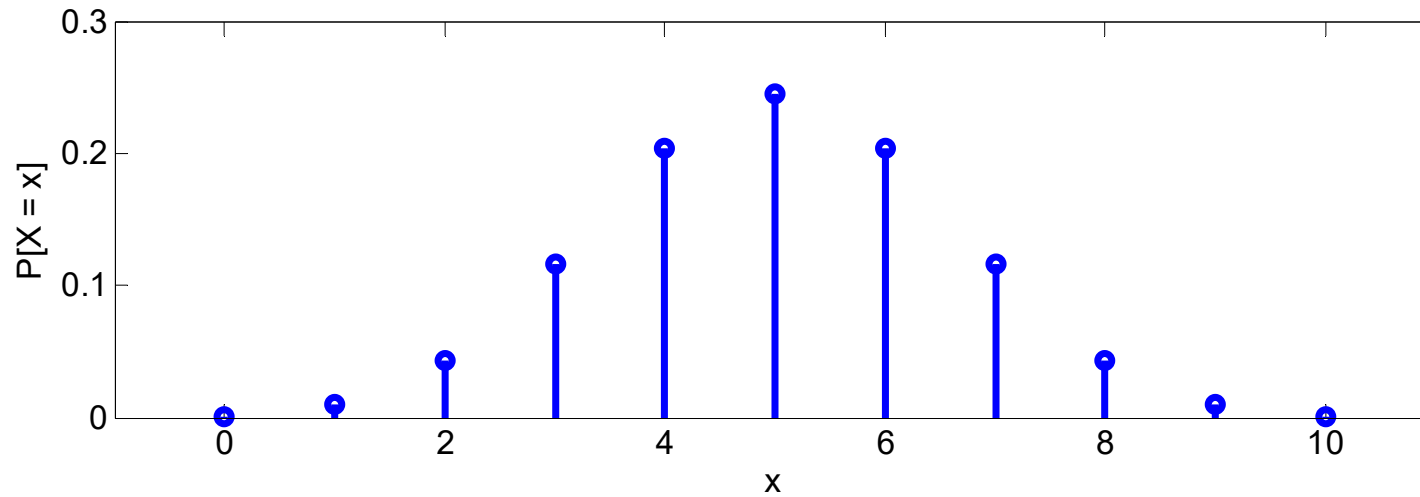
Binomialverteilung: $X \sim \text{Bin}(n, p)$

- X sei die Summe von n **unabhängigen** Bernoulli verteilten Zufallsvariablen mit **gleicher** Erfolgswahrscheinlichkeit p .
- D.h. X zählt die **Anzahl «Erfolge»** in n **unabhängigen «Versuchen»** (mit individueller Erfolgsw'keit p).
- Wertebereich: $W = \{0, 1, \dots, n\}$
- Wa'keitsfunktion:
$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x} \text{ für } x \in W \text{ (siehe Kapitel C.1 im Skript)}$$
- Erwartungswert und Varianz:
$$E(X) = np$$
$$\text{Var}(X) = np(1 - p)$$

Binomialverteilung ($n = 10, p = 0.3$)



Binomialverteilung ($n = 10, p = 0.5$)



Binomialverteilung: Beispiel

- Sei $p = 0.15$ die W'keit, dass eine Betonprobe mangelhaft ist.
(«Erfolg» = mangelhaft)
- Sei X die Anzahl mangelhafter Proben von insgesamt 10 (unabhängigen) Proben.
- X ist also Binomial-verteilt: $X \sim \text{Bin}(n, p)$ mit $n = 10$ und $p = 0.15$.
- Damit können wir diverse Sachen berechnen.

Binomialverteilung: Beispiel

- Was ist die W'keit, dass von 10 Proben **keine** mangelhaft ist?

$$P(X = 0) = \binom{10}{0} p^0 (1 - p)^{10} = (1 - p)^{10} = 0.85^{10}$$

- Was ist die W'keit, dass von 10 Proben **mindestens eine** mangelhaft ist?

$$P(X \geq 1) = 1 - P(X = 0) = 1 - 0.85^{10}$$

- Was ist die W'keit, dass von 10 Proben **genau zwei** mangelhaft sind?

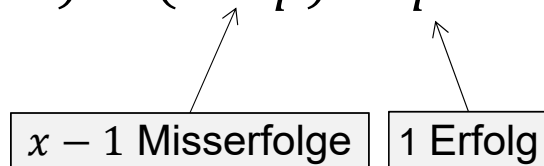
$$P(X = 2) = \binom{10}{2} p^2 (1 - p)^8 = 45 \cdot 0.15^2 \cdot 0.85^8$$

Binomialverteilung: Bemerkungen

- Die «Anzahl Versuche» n ist in der Regel aus dem Kontext vorgegeben.
- Die W'keit p ist dagegen ein **Parameter**, der in der Regel **unbekannt** ist (bis jetzt haben wir einfach entsprechende Annahmen getroffen).
- Typischerweise will man p aus Daten schätzen (siehe später).
- Bis auf Weiteres tun wir so, als ob wir p kennen würden.

Geometrische Verteilung: $X \sim \text{Geom}(s)$

- X sei die **Anzahl Wiederholungen** von unabhängigen Bernoulli(s) Experimenten **bis zum ersten Erfolg**.
- Beispiel: Werfe eine Münze so lange, bis zum **ersten Mal** Zahl erscheint und notiere die Anzahl Würfe.
- Wertebereich: $W = \{1, 2, 3, \dots\}$ (**unbeschränkt!**)
- Es gilt $P(X = x) = (1 - p)^{x-1}p$ für $x \in W$.



- Bemerkung: Wir brauchen jetzt keinen Binomialkoeffizient, weil die Misserfolge alle bevor dem Erfolg passieren müssen.

Geometrische Verteilung

- Erwartungswert und Varianz

$$E(X) = 1/p;$$

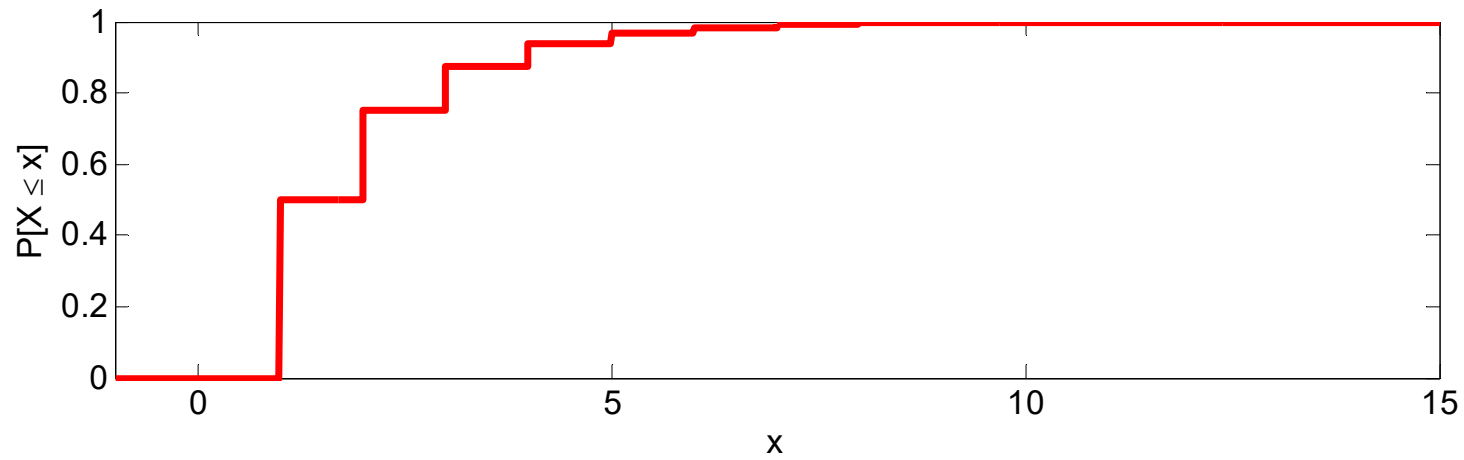
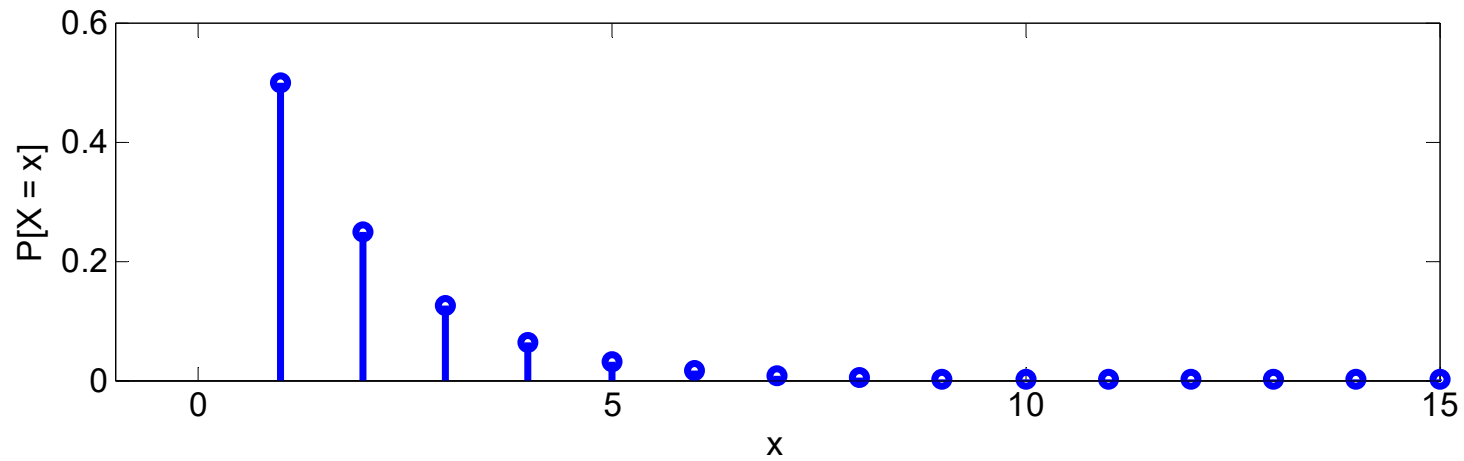
$$Var(X) = \frac{1-p}{p^2}$$

- Erwartungswert entspricht der **mittleren Wartezeit bis zum ersten «Erfolg»**, wird auch als **Wiederkehrperiode** bezeichnet.
- Die kumulative Verteilungsfunktion hat eine einfache Form:

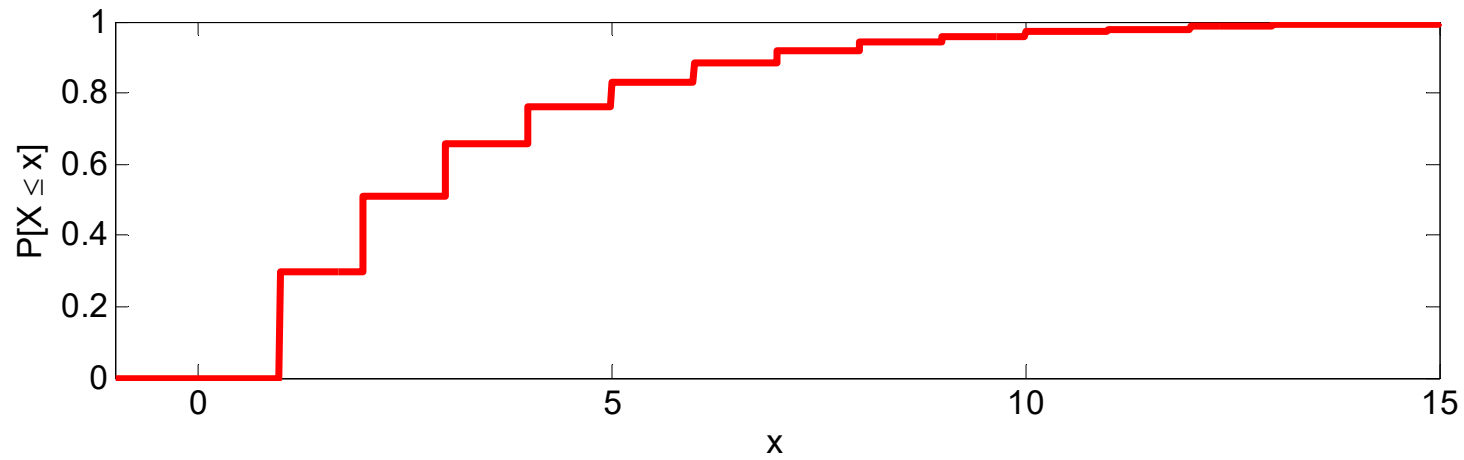
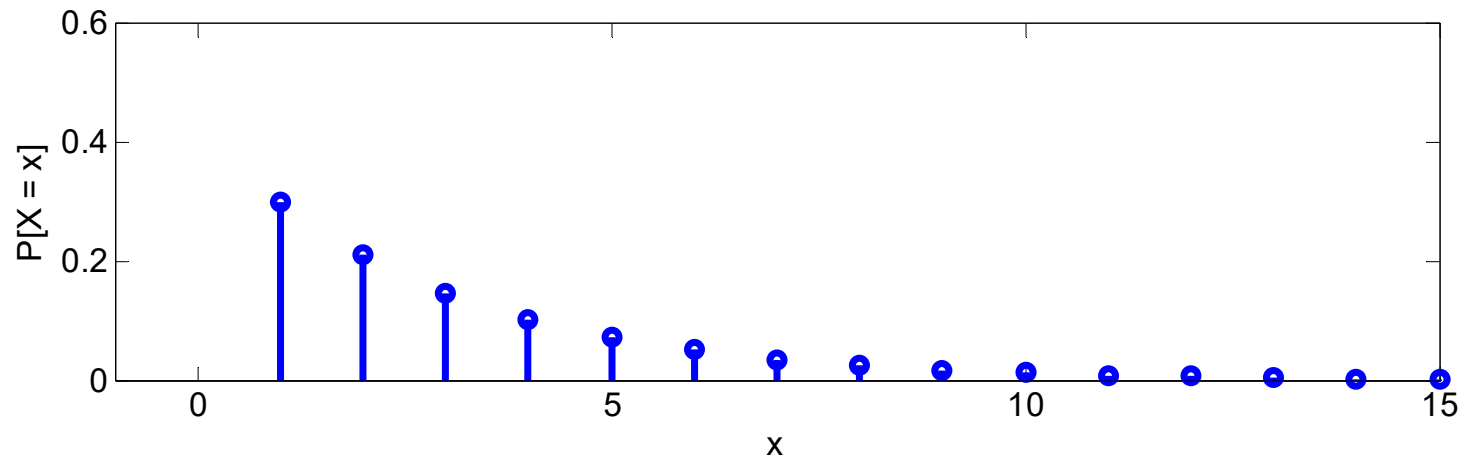
$$F(x) = \sum_{i=1}^x (1-p)^{i-1} p = 1 - (1-p)^x \text{ für } x \in W.$$

(geometrische Reihe)

Geometrische Verteilung ($p = 0.5$)



Geometrische Verteilung ($p = 0.3$)



Poisson-Verteilung: $X \sim \text{Pois}(\lambda)$, $\lambda > 0$

- Anwendung: Modellierung von (**unbeschränkten**) **Anzahlen**.

- Wertebereich: $W = \{0, 1, 2, \dots\}$ (**unbeschränkt**)

- Wa'keitsfunktion:

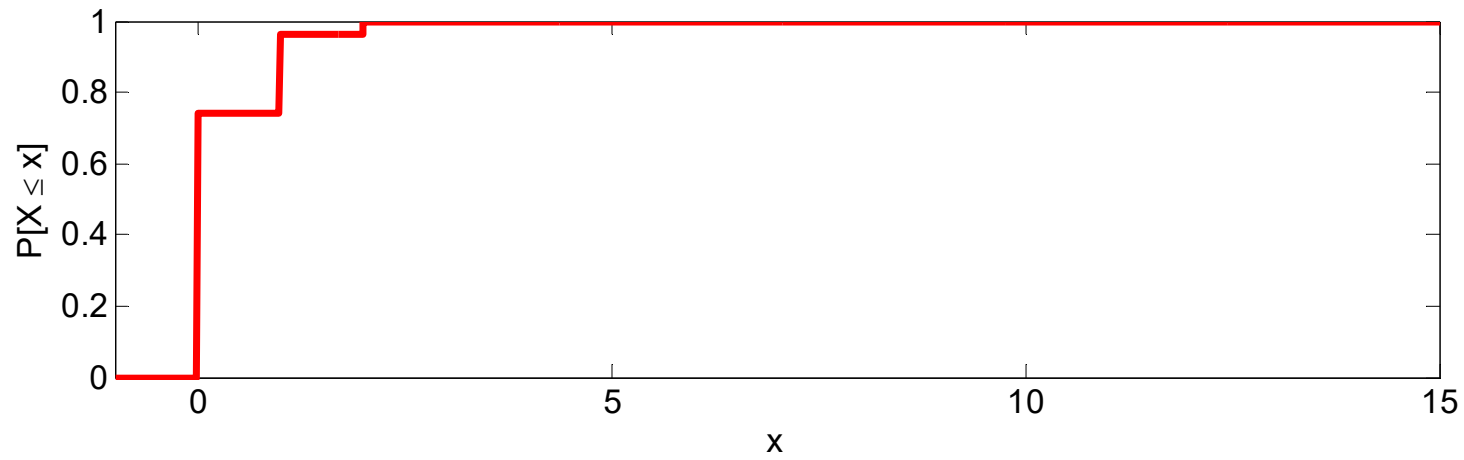
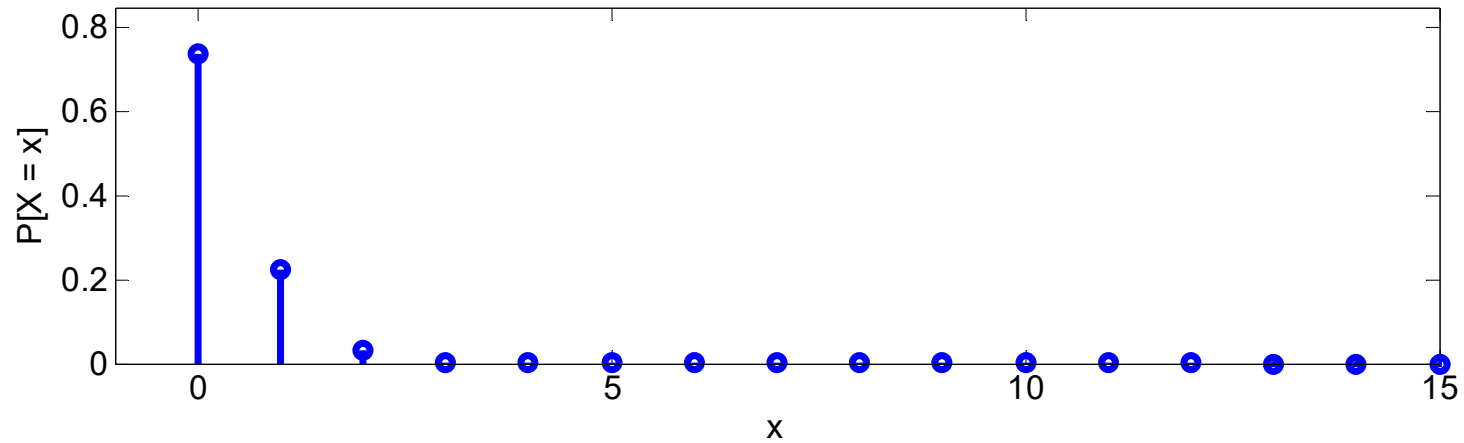
$$P(X = x) = e^{-\lambda} \frac{\lambda^x}{x!} \text{ für } x \in W.$$

- Erwartungswert / Varianz (hier identisch!)

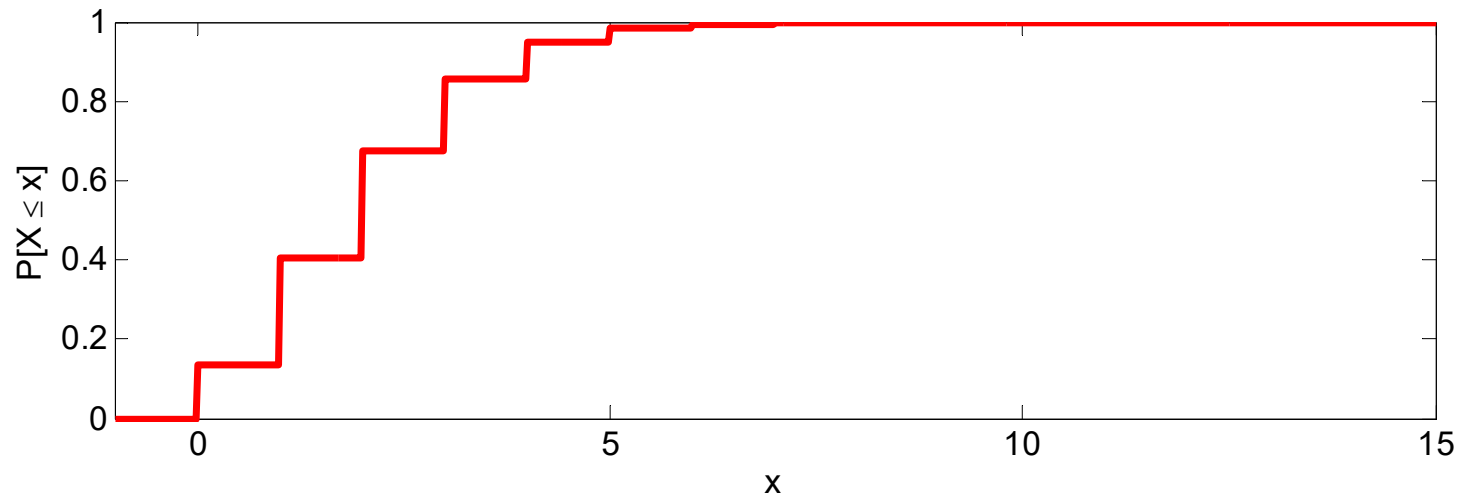
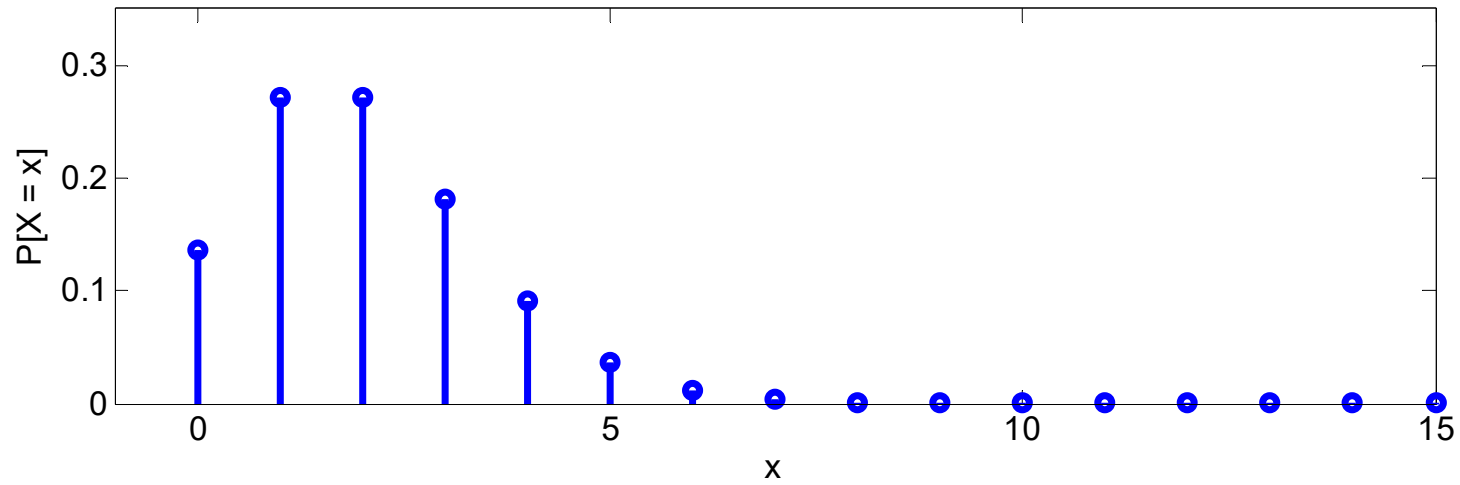
$$E(X) = \lambda$$

$$\text{Var}(X) = \lambda$$

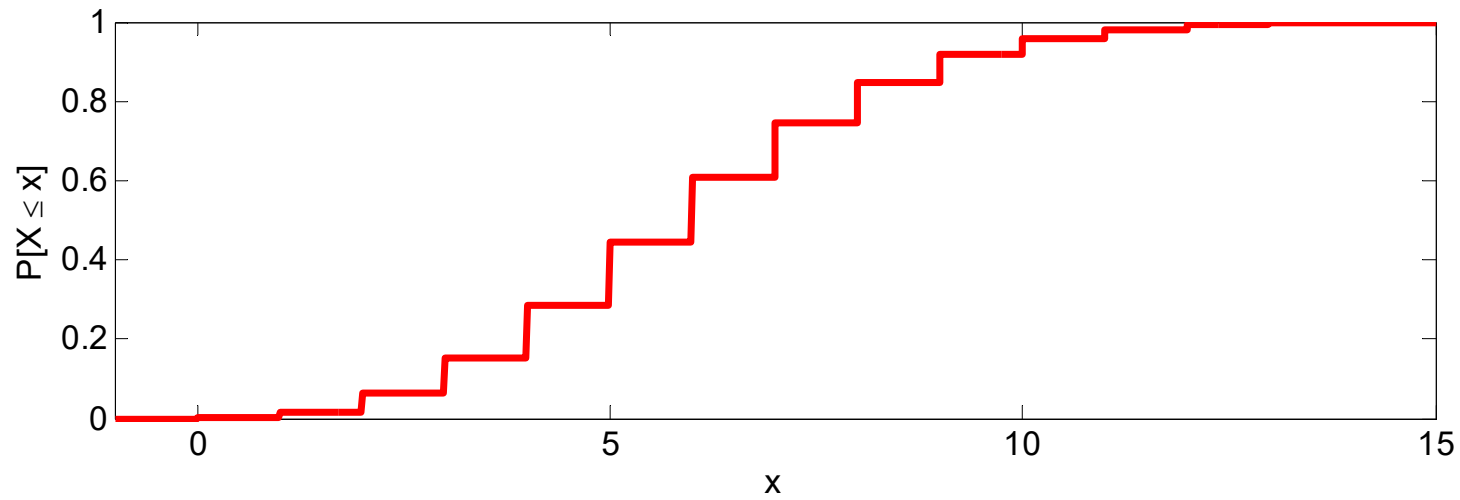
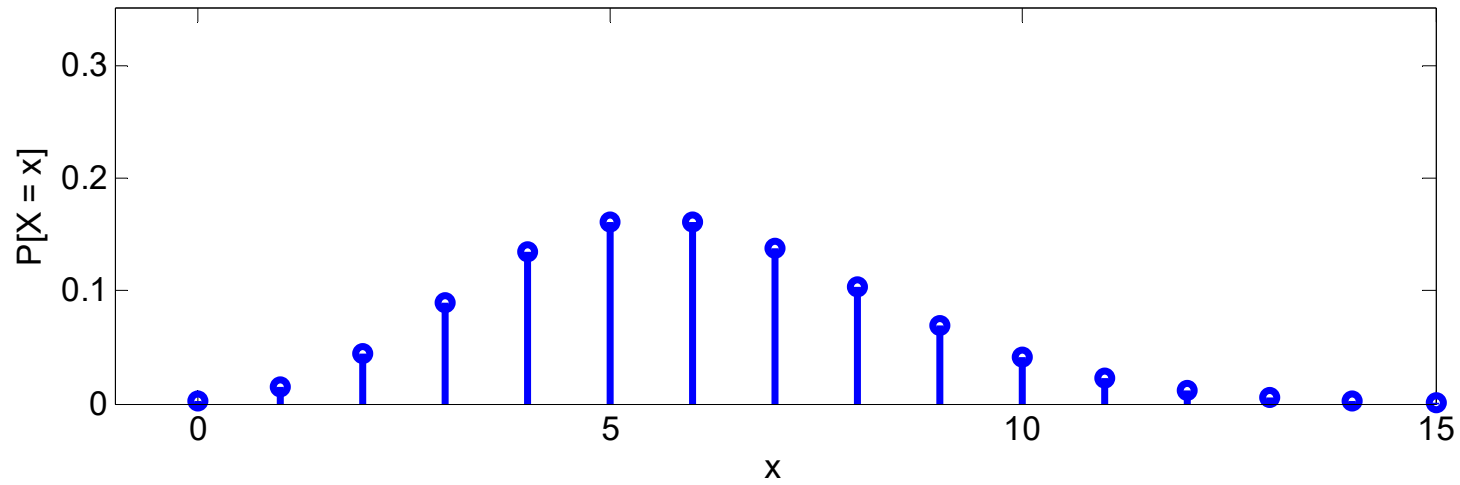
Poisson-Verteilung ($\lambda = 0.3$)



Poisson-Verteilung ($\lambda = 2$)



Poisson-Verteilung ($\lambda = 6$)



Poisson-Verteilung: Eigenschaften

- Es gilt: $Pois(\lambda) \approx Bin(n, p)$, für n gross, p klein und $np = \lambda$.
D.h. die Poisson-Verteilung kann interpretiert werden als Verteilung der Anzahl Ereignisse für seltene Ereignisse (p klein) bei vielen (n gross) unabhängigen Versuchen.
- Ist $X \sim Pois(\lambda_1)$, $Y \sim Pois(\lambda_2)$, **unabhängig**, dann ist

$$X + Y \sim Pois(\lambda_1 + \lambda_2)$$

Poisson-Verteilung: Beispiel

- Ein Callcenter wird im Schnitt alle 12 Sekunden angerufen. D.h. wir haben **im Schnitt 5 Anrufe pro Minute**.
- Die Anzahl Anrufe pro Minute modellieren wir mit der Zufallsvariablen X , für die wir eine Poissonverteilung annehmen (**viele** Leute könnten potentiell anrufen, jeder hat jedoch eine **tiefe W'keit** dies zu tun):
- Was ist die W'keit, dass innerhalb einer Minute **keine** Anrufe eingehen?

$$X \sim \text{Pois}(5); \quad (\text{weil } E(X) = 5)$$

$$P(X = 0) = e^{-\lambda} \frac{\lambda^0}{0!} = e^{-5} \approx 0.00674$$

Allgemeines Vorgehen

Fragestellung (betreffend unsicherem Phänomen)

Bsp: Erwartete Anzahl Überschwemmungen in den nächsten 10 Jahren?

Annahmen

Bsp: Jährliche W'keit sei $p = 0.1$, Ereignisse in verschiedenen Jahren seien unabhängig.

Modell

Bsp: Modelliere Anzahl Überschwemmungen mit Binomialverteilung.

Antwort (basierend auf Modell)

Bsp: Erwartete Anzahl, W'keit dass keine Überschwemmung, etc.