Diskrete Verteilungen

Bernoulli-Verteilung: $X \sim$ Bernoulli(p)

Symbol für «verteilt wie»

- «Experiment» mit **zwei** Ausgängen: «Erfolg» ($X = 1$) oder «Misserfolg» ($X = 0$). Die Erfolgswahrscheinlichkeit sei p.

- Wertebereich von X: $W = \{0,1\}$

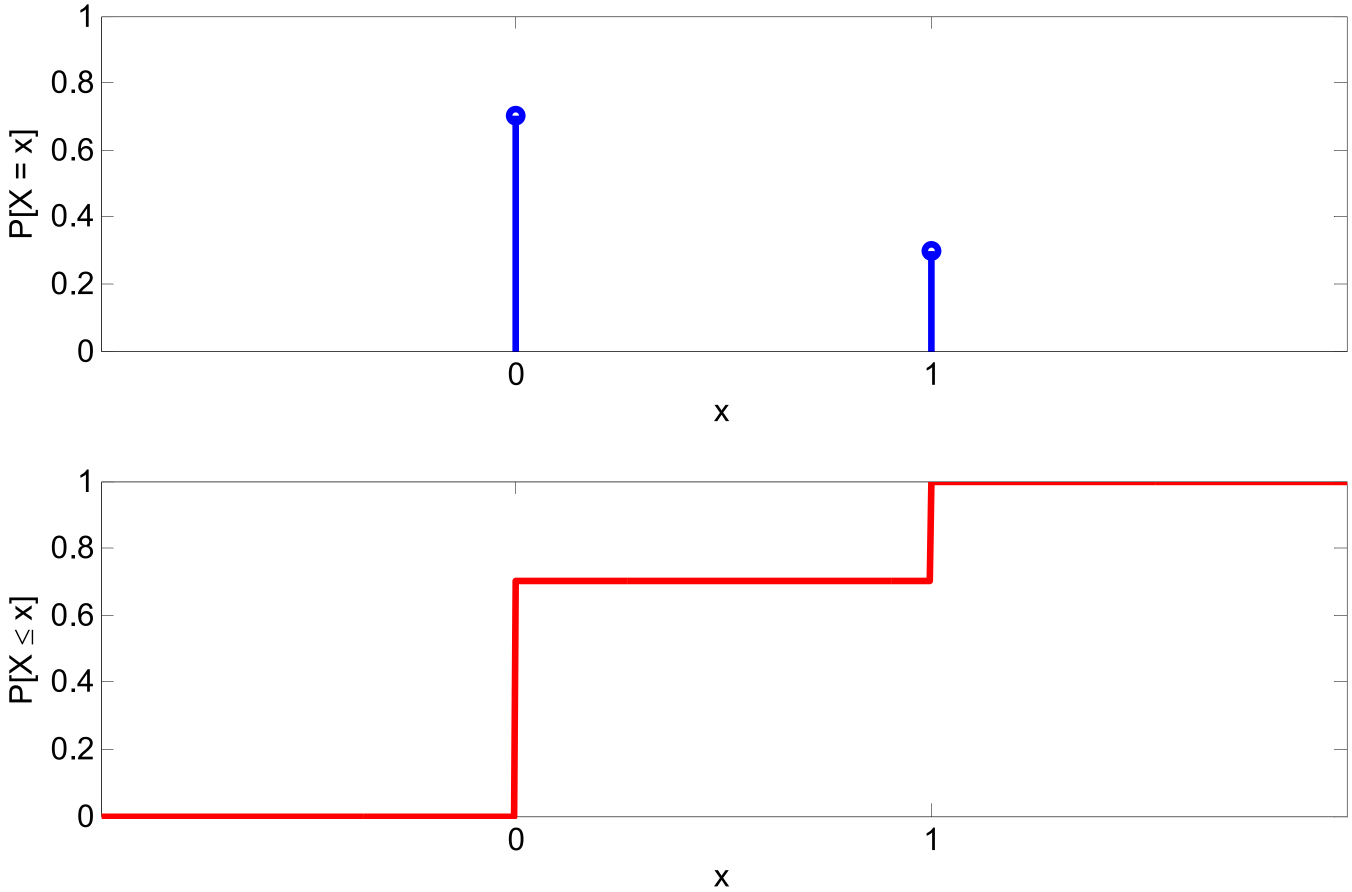
- D.h. wir haben

$$X = \begin{cases} 0 & \text{W'keit } 1 - p \\ 1 & \text{W'keit } p \end{cases}$$

- Erwartungswert / Varianz (nachrechnen!)

$$E[X] = p \qquad Var(X) = p(1-p)$$

Bernoulli-Verteilung $(p = 0.3)$



Bernoulli-Verteilung

«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).
- etc.

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

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

- X sei die Summe von n **unabhängigen** Bernoulli-verteilten Zufallsvariablen mit **gleicher** Erfolgswahrscheinlichkeit p.
- D.h. X zählt die **Anzahl der «Erfolge» in n unabhängigen «Versuchen»** (mit individueller Erfolgsw'keit p).
- Der Wertebereich ist daher $W = \{0, 1, \ldots, n\}$
- Es ist

$$P[X = x] = \binom{n}{x} p^x (1-p)^{n-x},\ x \in W$$

wobei

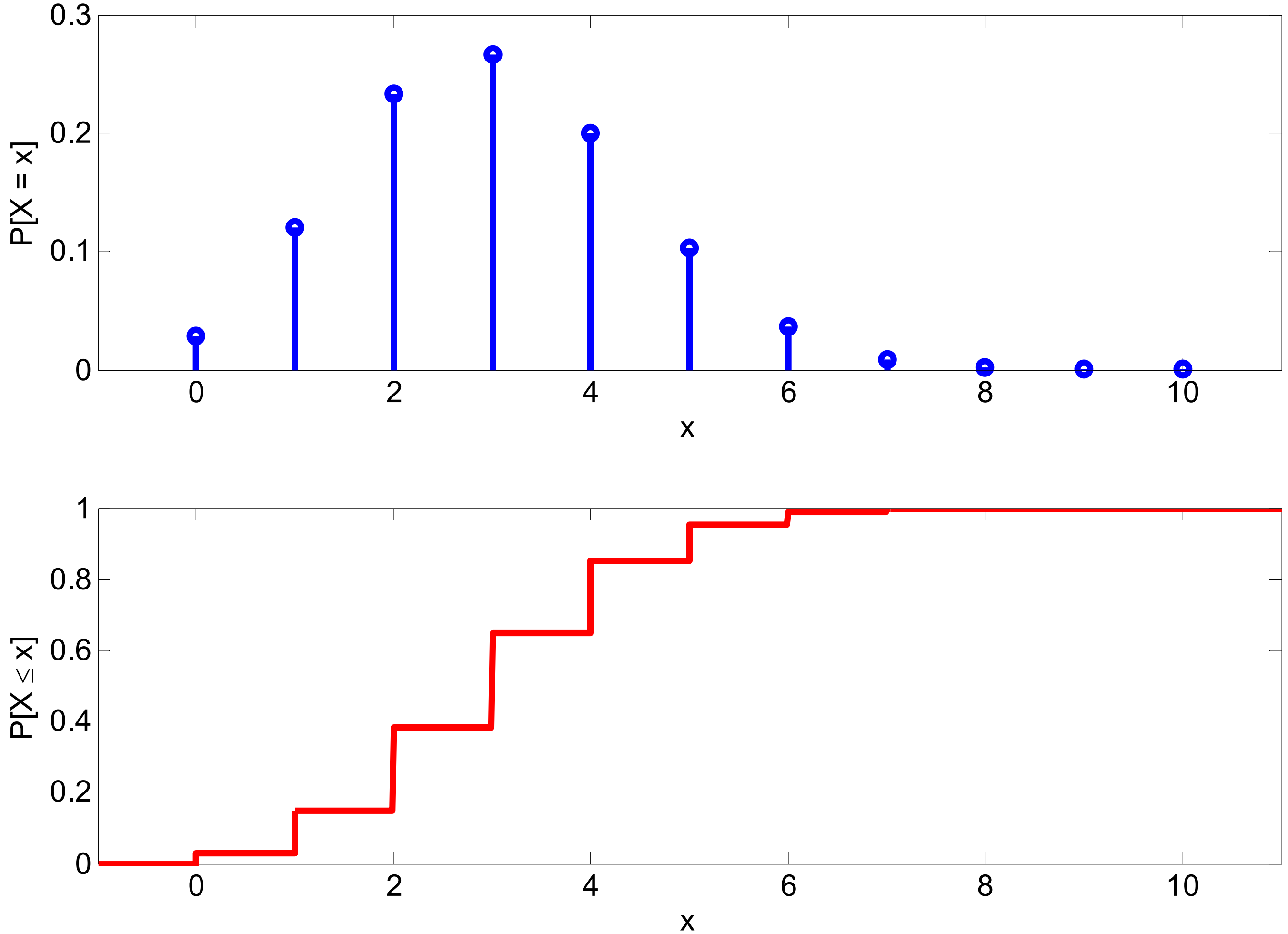
$$\binom{n}{x} = \frac{n!}{x!(n-x)!} = \frac{n(n-1)\cdots(n-x+1)}{x!}$$

Binomialverteilung

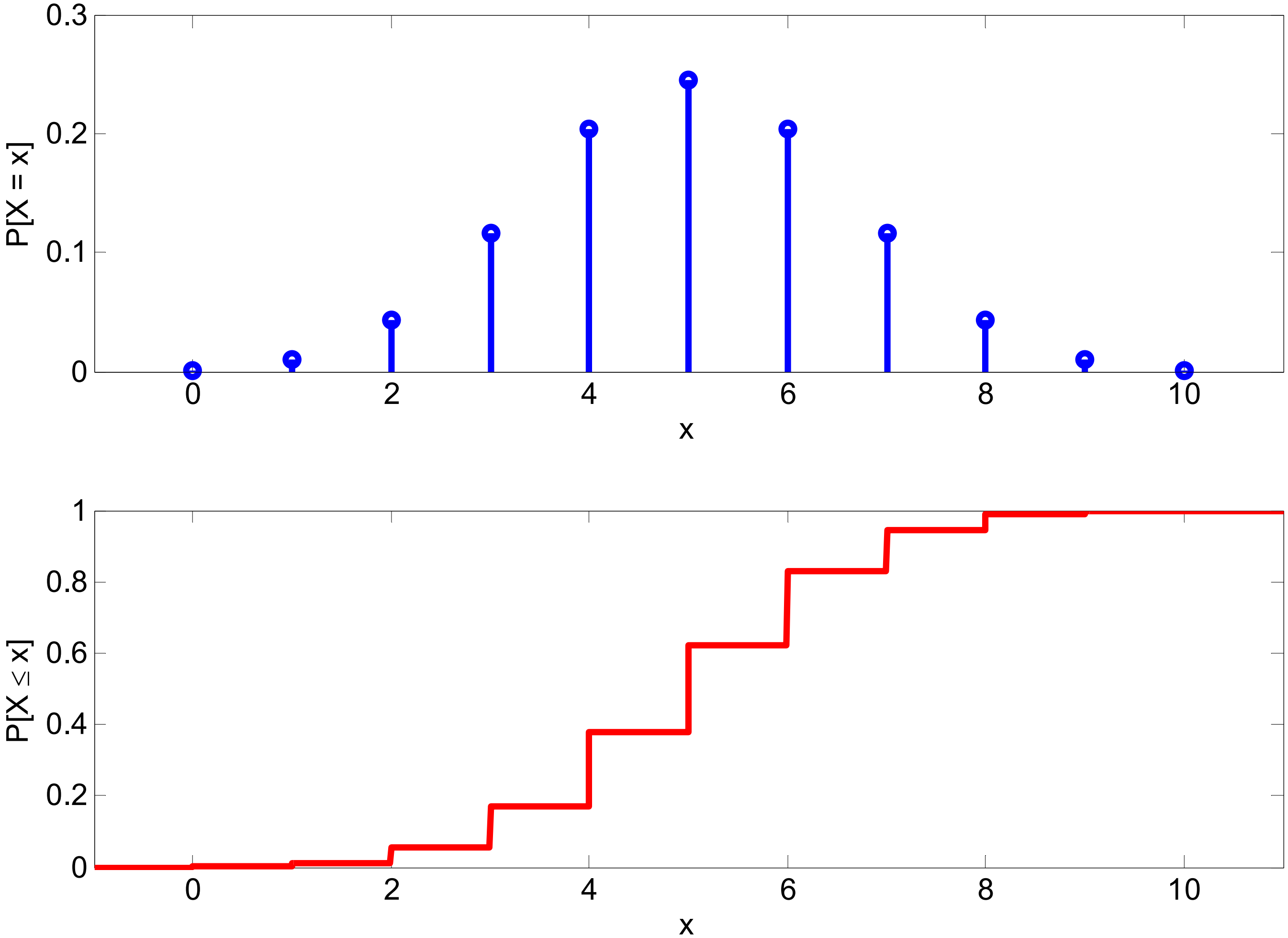
- $\binom{n}{x}$ heisst **Binomialkoeffizient** und ist die Anzahl der Möglichkeiten, aus n Objekten x auszuwählen, **ohne** Berücksichtigung der Reihenfolge.
- Bsp: Wähle aus 20 Studenten 3 aus:
  $\binom{20}{3}$ verschiedene Möglichkeiten eine 3-er Gruppe zu bilden
- Erwartungswert / Varianz

$$E[X] = np \qquad Var(X) = np(1-p)$$

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



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



Binomialverteilung: Beispiel

- p sei die W'keit, dass eine Betonprobe den erforderlichen Anforderungen **nicht** entspricht, z.B. $p = 0.05$.

- X sei die Anzahl «mangelhafter» Proben von insgesamt 10 (unabhängigen) Proben.

- X ist also Binomial-verteilt: $X \sim Bin(n, p)$ mit $n = 10$ und $p = 0.05$.

- Damit können wir diverse Sachen berechnen.

Binomialverteilung: Beispiel

- Was ist die W'keit, dass von 10 Proben **alle** den Anforderungen entsprechen?

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

- Was ist die W'keit, dass von 10 Proben **mindestens eine** den Anforderungen **nicht** entspricht?

$$P[X \geq 1] = 1 - P[X = 0] = 1 - (1-p)^{10}$$

- Was ist die W'keit, dass von 10 Proben **genau eine** den Anforderungen **nicht** entspricht?

$$P[X = 1] = \binom{10}{1} p^1 (1-p)^9 = 10 \cdot p^1 (1-p)^9$$

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: $X \sim$ Geom(p)

- X sei die **Anzahl der Wiederholungen** von unabhängigen Bernoulli(p)-Experimenten **bis zum ersten Erfolg**.

- Bsp: Werfe Münze so lange, bis zum **ersten Mal** Zahl erscheint. Notiere die Anzahl der Würfe.

- Wertebereich: $W = \{1, 2, 3, \ldots\}$ (**unbeschränkt!**)

- Es gilt

$$P[X = x] = \underbrace{(1-p)^{x-1}}_{x-1 \text{ Misserfolge}} \underbrace{p}_{1 \text{ Erfolg}}, \; x \in W.$$

Geometrische Verteilung

- Kumulative Verteilungsfunktion

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

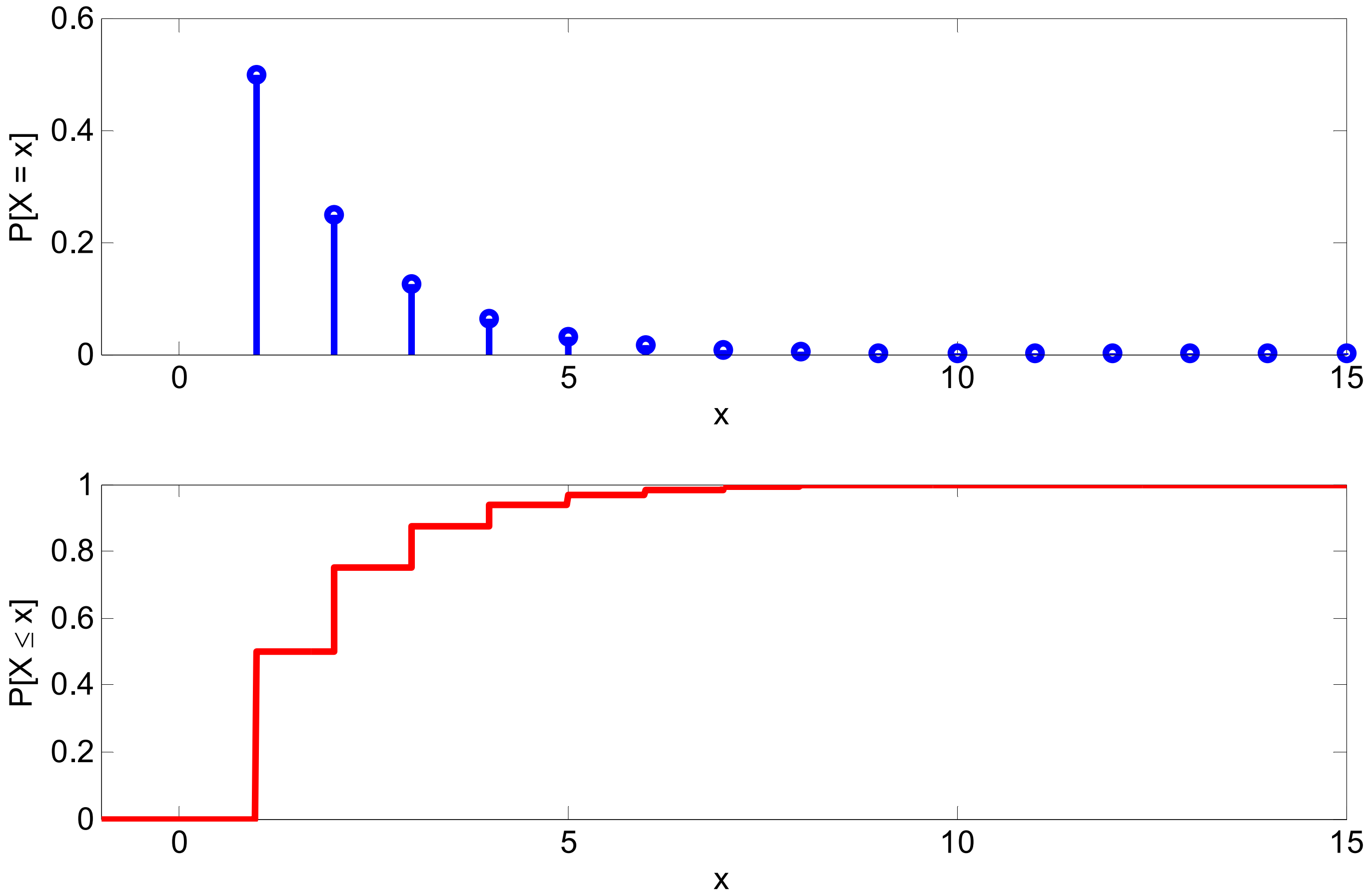
(dazwischen ist F **konstant**)

- Erwartungswert / Varianz

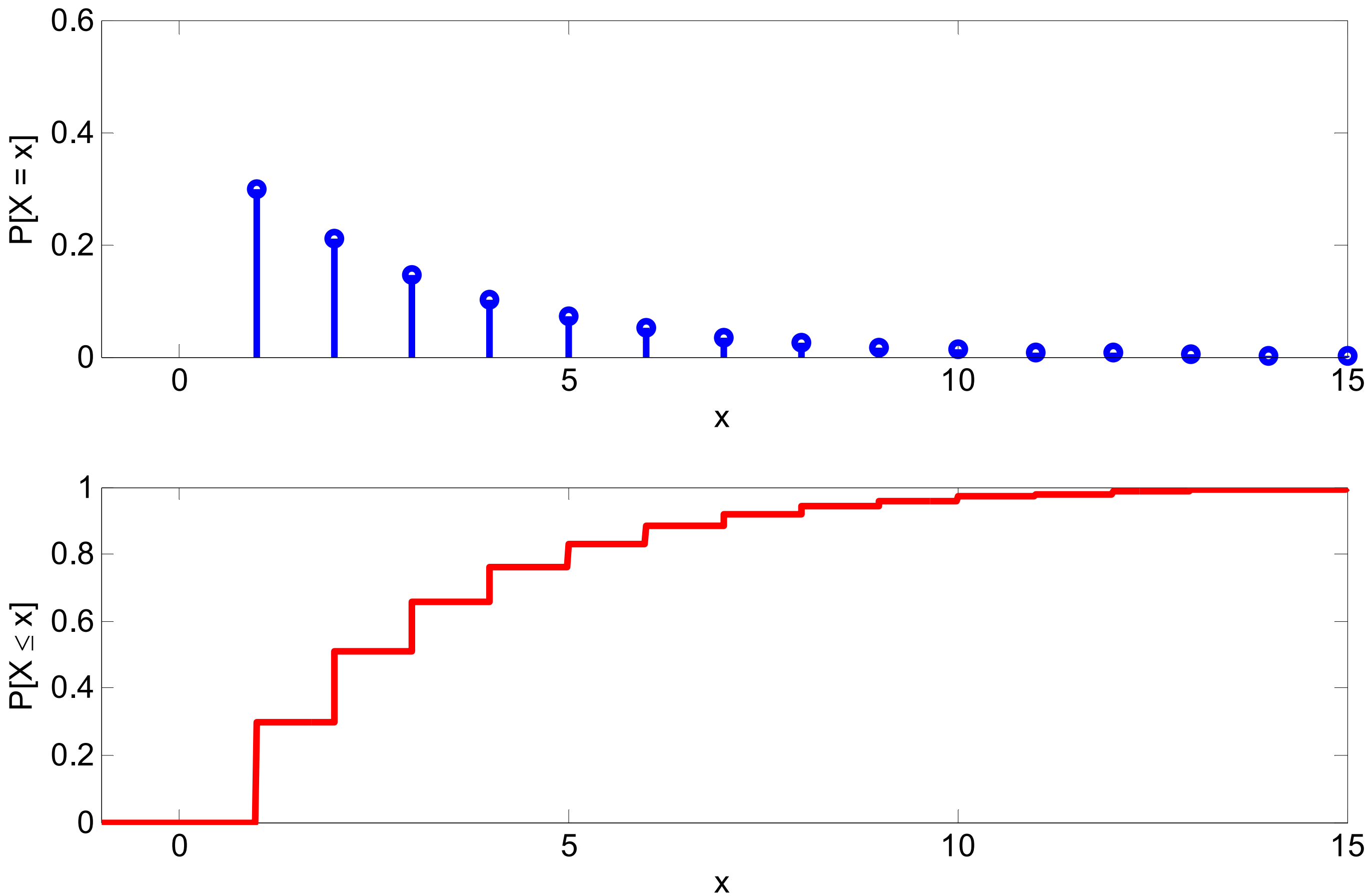
$$E[X] = \frac{1}{p} \qquad Var(X) = \frac{1-p}{p^2}$$

- Erwartungswert entspricht der **mittleren Wartezeit bis zum ersten «Erfolg»**, wird auch als **Wiederkehrperiode** bezeichnet.

Geometrische Verteilung $(p = 0.5)$



Geometrische Verteilung $(p = 0.3)$



Poisson-Verteilung: $X \sim$ Pois(λ), λ > 0

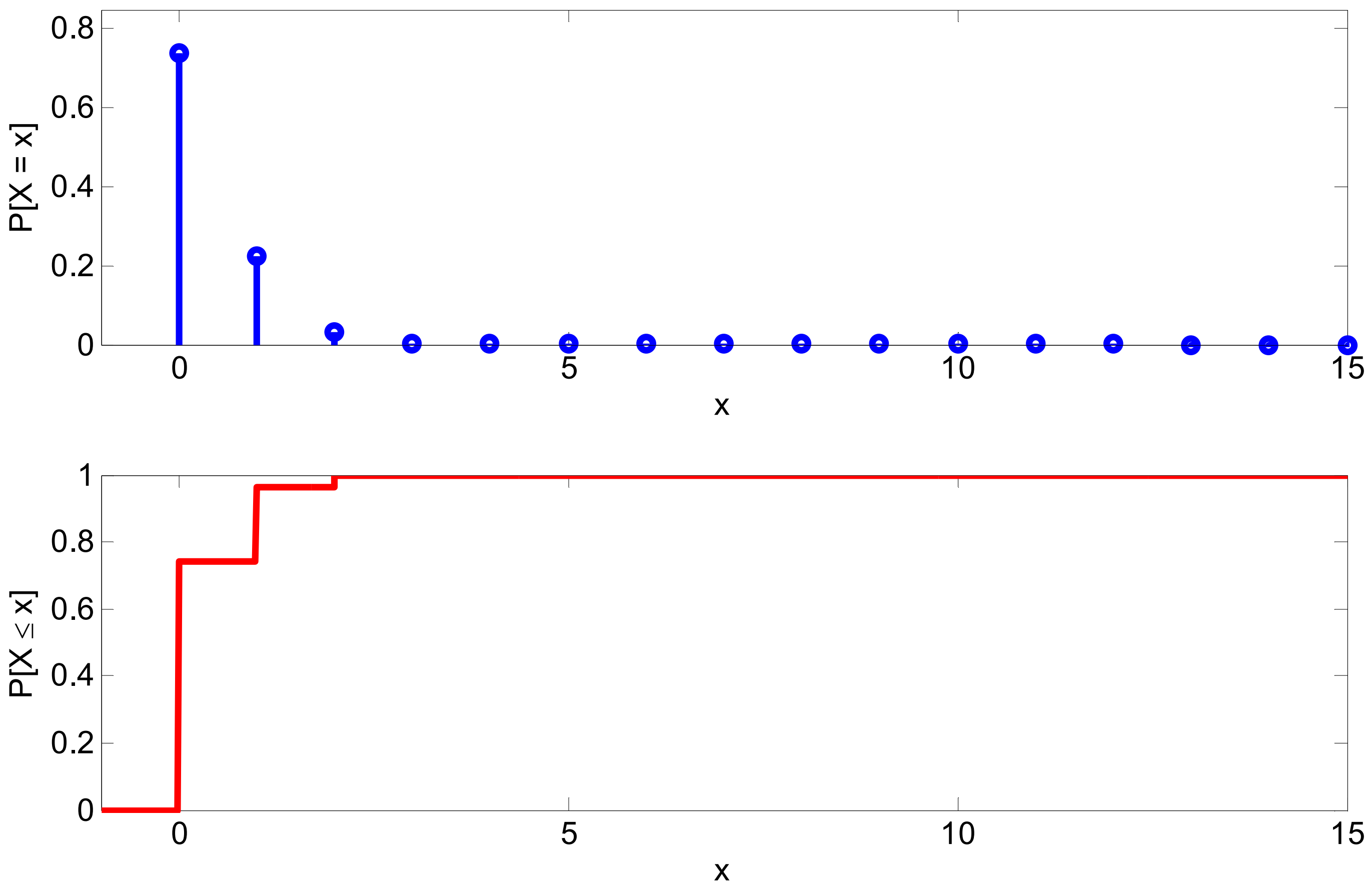
- Anwendung: Modellierung von (**unbeschränkten**) **Anzahlen**.
- Wertebereich: $W = \{0, 1, 2, ...\}$ (**unbeschränkt**)
- Es ist

$$P[X = x] = e^{-\lambda} \frac{\lambda^x}{x!}, \ x \in W$$

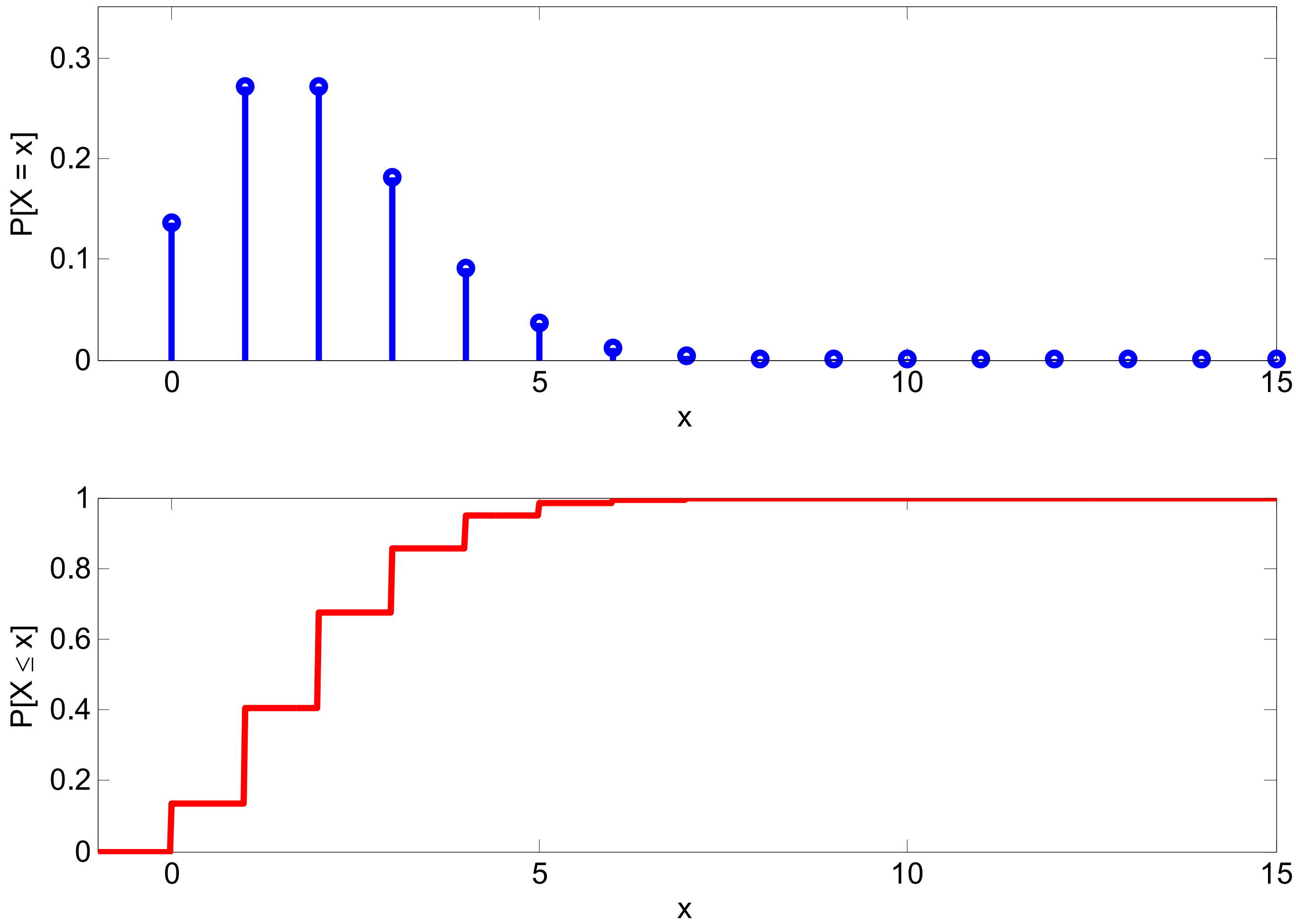
- Erwartungswert / Varianz (hier identisch!)

$$E[X] = \lambda \qquad Var(X) = \lambda$$

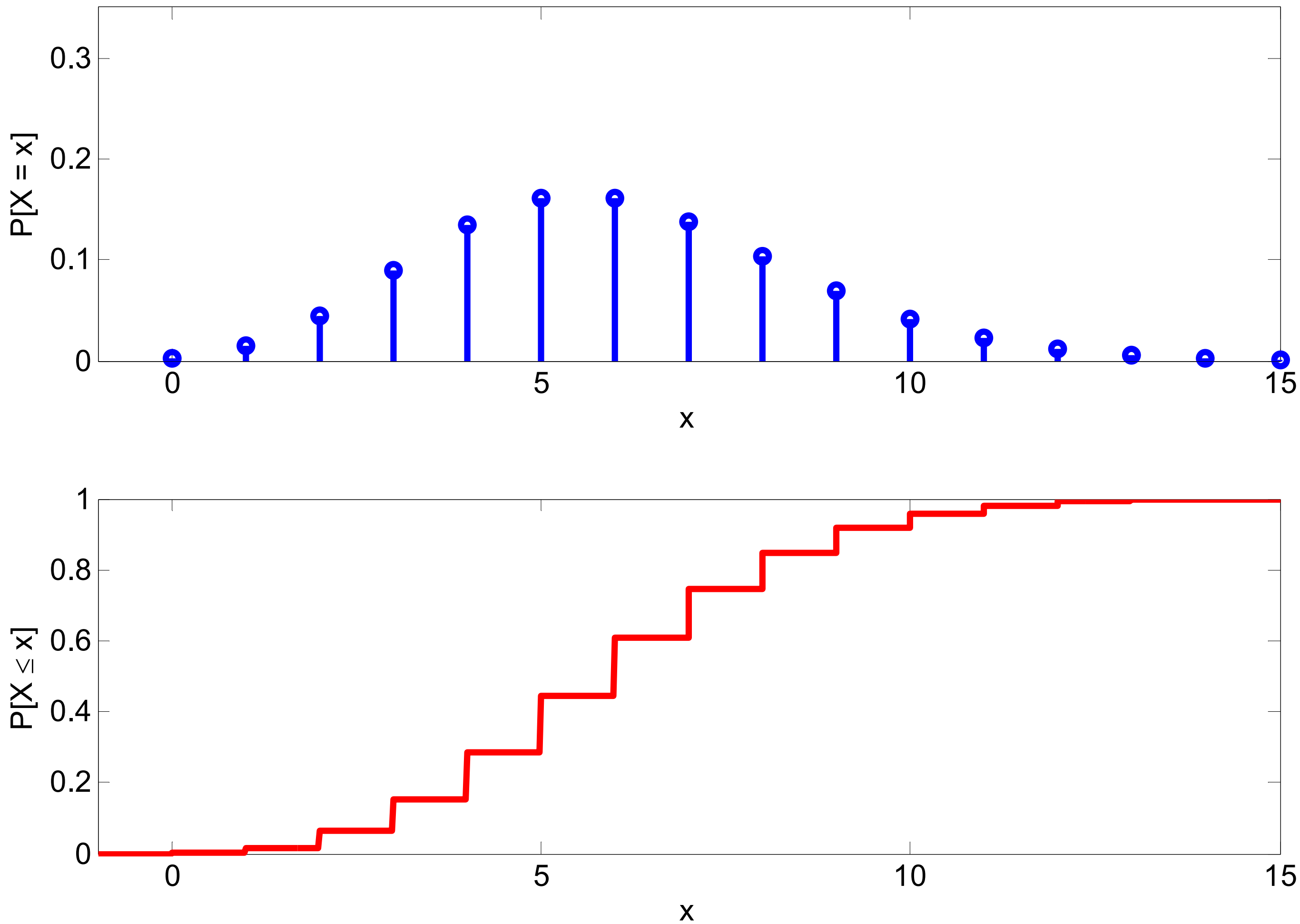
Poisson-Verteilung $(\lambda = 0.3)$



Poisson-Verteilung $(\lambda = 2)$



Poisson-Verteilung $(\lambda = 6)$



Poisson-Verteilung: Eigenschaften

- Es gilt: Pois(λ) $\approx$ Bin(n, p), für n gross, p klein und $np = \lambda$.

  D.h. die Poisson-Verteilung kann interpretiert werden als Verteilung für seltene Ereignisse (p klein) bei vielen (n gross) unabhängigen Versuchen.

- Ist X ~ Pois(λ_1), Y ~ Pois(λ_2), **unabhängig**, dann ist

$$X + Y \sim \mathrm{Pois}(\lambda_1 + \lambda_2)$$

- Die Anzahl der Ausfälle eines Systems in einem Zeitintervall der Länge t kann z.B. als Pois(λt) modelliert werden.

- Erweiterung: Poissonprozess (später).

Poisson-Verteilung: Beispiel

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

$$X \sim \mathrm{Pois}(\lambda),\ \lambda = 5 \quad (E[X] = \lambda = 5)$$

$$P[X = 0] = e^{-\lambda}\frac{\lambda^0}{0!} = e^{-\lambda} = e^{-5} = 0.00674$$

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 der Überschwemmungen mit Binomialverteilung.

Antwort (**basierend auf Modell**)

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