

Induktive Statistik für Soziologinnen und Soziologen (Mariana Nold)

Thema: Zufallsvorgang, Inferenzschluss und Wahrscheinlichkeitsverteilungen

Abgabe: keine Abgabe, wird in der Übung besprochen

Wichtige Definitionen:

1. Zufallsvorgang
2. Zufallsvariable
3. Einfache Zufallsstichprobe
4. Wahrscheinlichkeitsverteilung
5. Binäre Variable und Bernoulliverteilung
6. Binomialverteilung

1. Relative Häufigkeit beim Münzwurf

In der Vorlesung hatten wir die Wahrscheinlichkeitsfunktion der Binomialverteilung für den 10-fachen Münzwurf kennen gelernt. Die entsprechende Grafik ist gegeben durch Abbildung 1. Die Tabelle 1 enthält die entsprechenden Werte in Zahlen.

Diese Wahrscheinlichkeitsfunktion lässt sich auch in Form einer Tabelle angeben

Anzahl Treffer y	0	1	2	3	4	5	6	7	8	9	10
W'keit $\mathbb{P}(Y = y)$	0.001	0.010	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.010	0.001

Tabelle 1: Wahrscheinlichkeitsfunktion der $Bin(n = 10, p = 0.5)$ -Verteilung

- (a) Nehmen Sie an, dass 200 Studierende an dem Experiment 10-facher Münzwurf teilgenommen haben. Jede Person hat 10 Mal eine Münze geworfen und die Anzahl der Treffer (Münze zeigt Zahl) auf einem Zettel notiert. Warum handelt es sich um einen Zufallsvorgang?
- (b) Nutzen Sie die Zufallsvariable Y , um den Zufallsvorgang formal zu beschreiben.
- (c) Sie wählen zufällig eine Person von diesen 200 Studierenden aus und fragen sie, ob sie auf ihrem Zettel die Zahl 4 notiert hat. Wie hoch ist die Wahrscheinlichkeit, dass die Person mit “ja” antwortet?

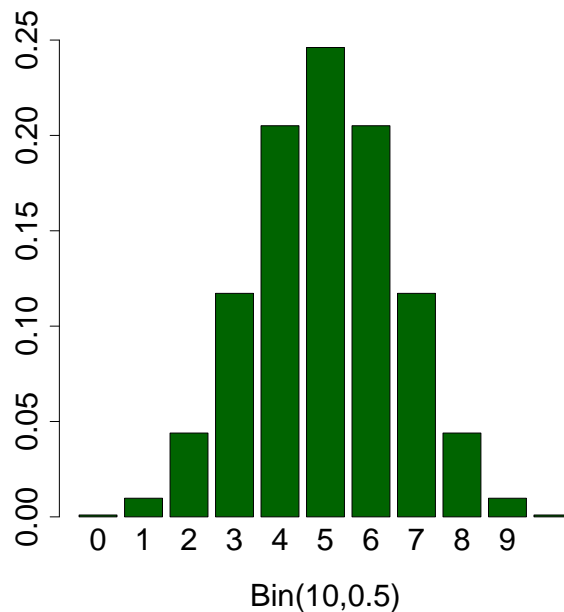


Abbildung 1: Wahrscheinlichkeitsfunktion der Binomialverteilung $p = 0.5$ und $n = 10$

- (d) Nach welcher Zahl y müssen Sie fragen, damit ihre Wahrscheinlichkeit eine Ja-Antwort zu bekommen maximal ist?
- (e) Beruhend auf den Überlegungen in Teilaufgabe d): Wo liegt der Modus der Wahrscheinlichkeitsfunktion der $Bin(n = 10, p = 0.5)$ -Verteilung?
- (f) Wie hoch ist die Wahrscheinlichkeit, dass eine zufällig ausgewählte Person tatsächlich 10 Mal in Folge einen Treffer erzielt hat?
- (g) Sie bitten alle Studierenden, die genau einen Treffer erzielt haben, sich zu melden. Wie viele Meldungen erwarten Sie beruhend auf den Angaben in Tabelle 1.
- (h) Wie viele Meldungen erwarten Sie beruhend auf den Angaben in Tabelle 1, wenn Sie alle Studierenden bitten, sich zu melden, die genau sechs Treffer erzielt haben?

2. **Erwartungswert und Mittelwert beim Münzwurf** In dieser Aufgabe betrachten wir nochmal den Zufallsvorgang aus Aufgabe 1. 200 Studierende werfen je 10 Mal eine Münze. Es sei

$$Y_i := \text{Anzahl der Treffer bei 10 Münzwürfen der Person } i \in \{0, \dots, 200\}$$

Endsprechen bezeichnet y_i die Anzahl der Treffer die Person i erzielt hat. So beinhaltet z. B. $y_3 = 8$ die Information, dass die dritte Person genau acht Treffer erzielt hat.

Sie kennen aus dem letzten Semester das arithmetische Mittel \bar{y} . Es wird berechnet als

$$\bar{y} = \sum_{i=1}^{n=200} y_i.$$

Die Abbildung 2 zeigt die relative Häufigkeitsverteilung die bei der Durchführung des Zufallsvorgangs beobachtet wurde.

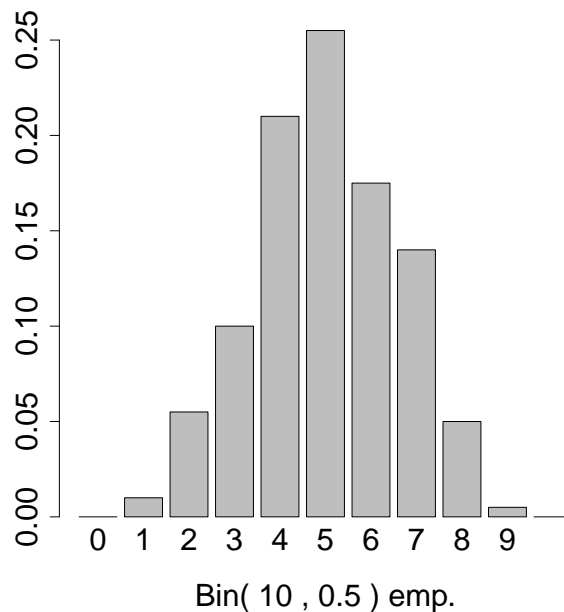


Abbildung 2: Wahrscheinlichkeitsfunktion der Binomialverteilung $p = 0.5$ und $n = 10$

Anzahl Treffer y	0	1	2	3	4	5	6	7	8	9	10
rel. H'keit	0.000	0.000	0.020	0.090	0.215	0.280	0.235	0.130	0.015	0.015	0.000
abs. H'keit	0	0	4	18	43	56	47	26	3	3	0

Tabelle 2: rel. Häufigkeit die beim Zufallsvorgang 10-facher Münzwurf von 200 Personen beobachtet wurde.

- Warum sind Abbildung 1 und Abbildung 2 ähnlich aber nicht gleich? Worin ähneln die beiden Abbildungen sich und wo sehen Sie Unterschiede? Zur besseren Vergleichbarkeit enthält Tabelle 2 die relativen Häufigkeiten, die in Abbildung 2 dargestellt sind.
- Vergleichen Sie die erwartete Anzahl der Meldungen aus Aufgabe 1 g) und 1 h) mit den tatsächlich beobachteten Häufigkeiten.
- Welcher Zusammenhang besteht zwischen der absoluten und der relativen Häufigkeit?
- Nutzen Sie die absolute Häufigkeit aus Tabelle 2 um den Mittelwert \bar{y} zu berechnen und interpretieren Sie diesen Wert.
- Wie kann das arithmetische Mittel \bar{y} beruhend auf den relativen Häufigkeiten berechnet werden?
- Nutzen Sie den gleichen Rechenansatz wie in Teilaufgabe e) um den Erwartungswert der $Bin(n = 10, p = 0.5)$ -Verteilung zu berechnen.

3. Analyse der Daten mit STATA Lesen Sie den Datensatz `munze.dta` ein.

- (a) Nutzen Sie den Befehl `. summarize` um eine Zusammenfassung der Daten zu erhalten.
- (b) Nutzen Sie den Befehl `. tab AnzTreffer` um eine Tabelle der Daten zu erhalten.
- (c) Lesen Sie aus der in Teilaufgabe e) erzeugten Tabelle ab, wie hoch der Anteil der Studierenden ist, der höchstens sechs Treffer erzielt hat.
- (d) Nutzen Sie den Befehl `. graph box AnzTreffer` um einen Boxplot zu erstellen und beschreiben Sie diesen Boxplot in Worten.