# Vorlesung 11a

# Bayes'sche Anteilschätzung und die Pólya-Urne

## Teil 2

## Die bedingte Verteilung der Erfolgswahrscheinlichkeit

(vgl. Buch S. 93)

Wieder betrachten wir die Situation als Teil 1:
In der ersten Stufe wird
die Erfolgswahrscheinlichkeit $P$ uniform aus $[0, 1]$ gewählt.
In der zweiten Stufe wird, gegeben $\{P = p\}$,
ein wiederholter $p$-Münzwurf $(Z_1, Z_2, \ldots)$ durchgeführt.
Sei $K_n$ die Anzahl der Erfoge in den ersten $n$ Versuchen.

Wenden wir uns jetzt der Frage 2 aus Teil 1 zu:

**Was ist die bedingte Verteilung von $P$,**

**gegeben $\{K_n = k\}$ ?**

Man spricht hier auch von der à posteriori Verteilung von $P$
(also quasi der Aktualisierung der Verteilung von $P$)
nach Beobachten von $K_n$.

Dazu ein Simulationsexperiment für $n = 2$:

Für 3000 unabhängig und uniform aus $[0, 1]$ gewählte $p$

wird jeweils ein 2-maliger $p$-Münzwurf durchgeführt.

Jeder der 3000 Punkte

gibt eine gemeinsame Realisierung

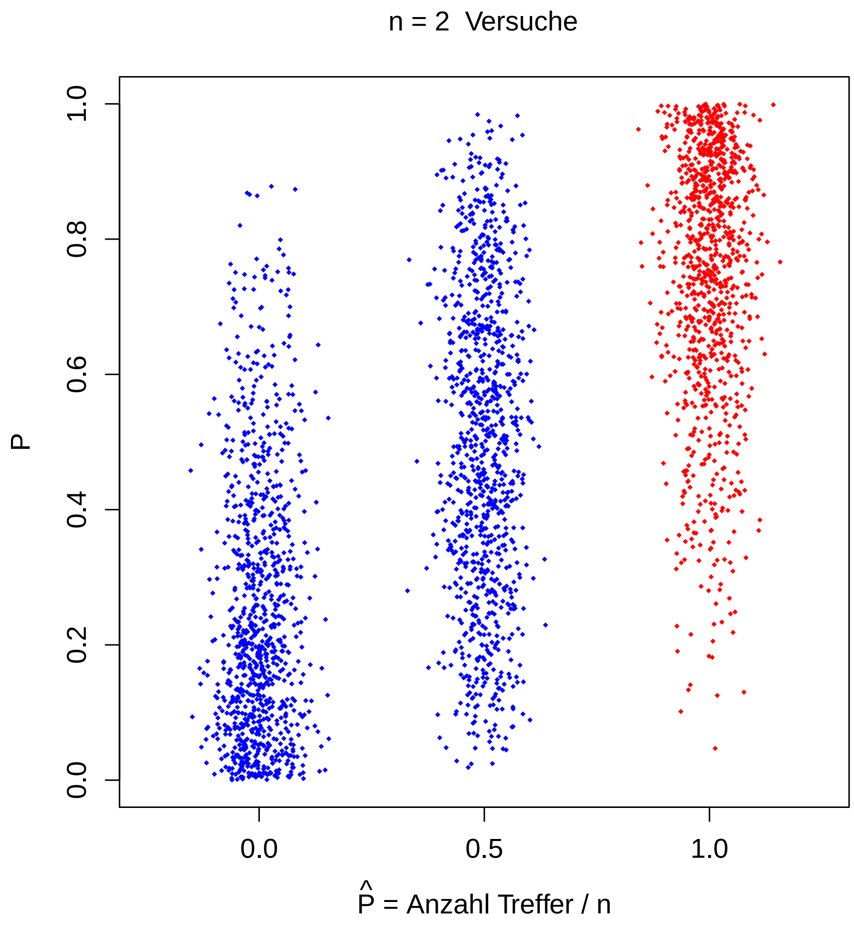
von $P$ und $\widehat{P} := K_2/2$.

Der Wertebereich von $P$ ist $[0, 1]$, der von $\widehat{P}$ ist $\{0, \frac{1}{2}, 1\}$.

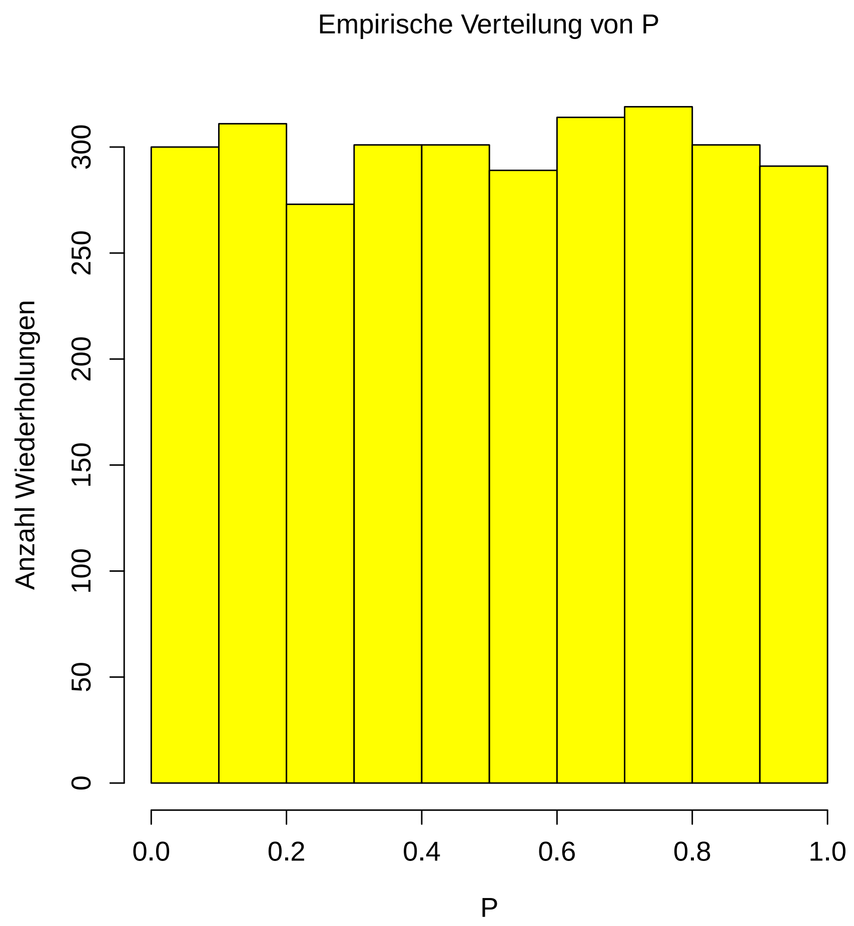
In der folgenden Abbildung sehen wir die 3000 Punkte,
mit $P$ vertikal und $\widehat{P}$ horizontal abgetragen.

Das horizontale Jittern (Verwackeln) von $\widehat{P}$
dient dabei (nur) dem Sichtbarmachen der Punkte.

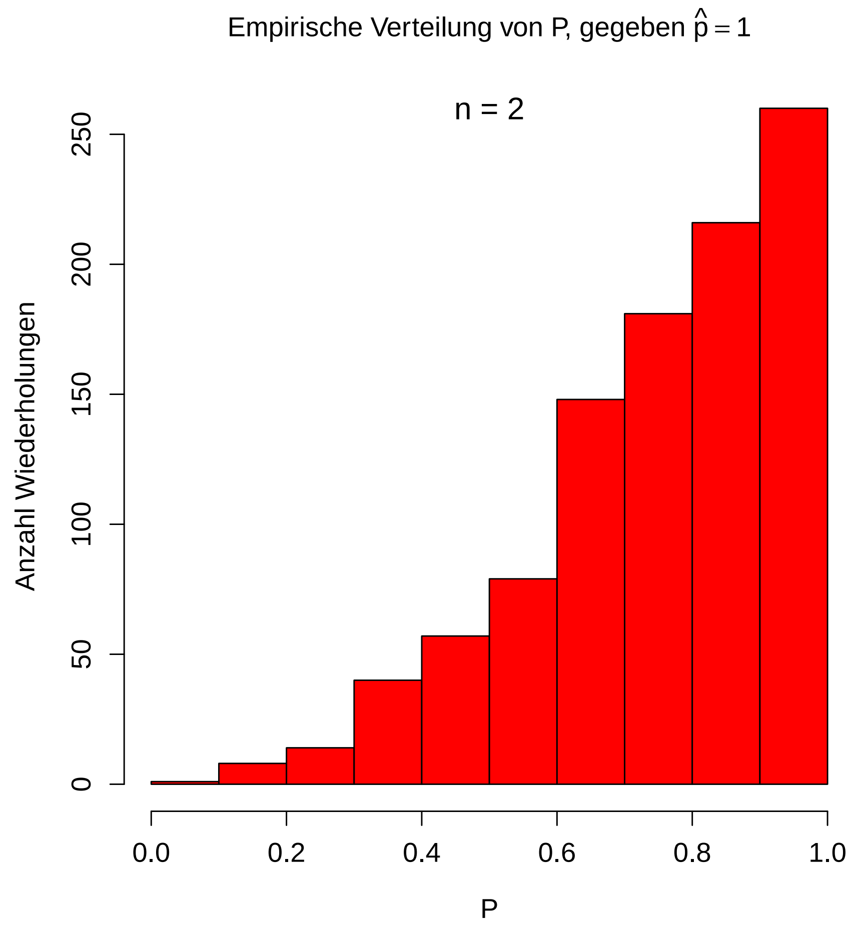
Gemeinsame Verteilung von $(P, K_2/2)$:



# A-Priori Verteilung von $P$



# Bedingte Verteilung von $P$ gegeben $\{K_2 = 2\}$



Jetzt für $n = 10$:

Für 11000 uniform aus $[0, 1]$ gewählte $p$

wird jeweils ein 10-maliger $p$-Münzwurf durchgeführt.

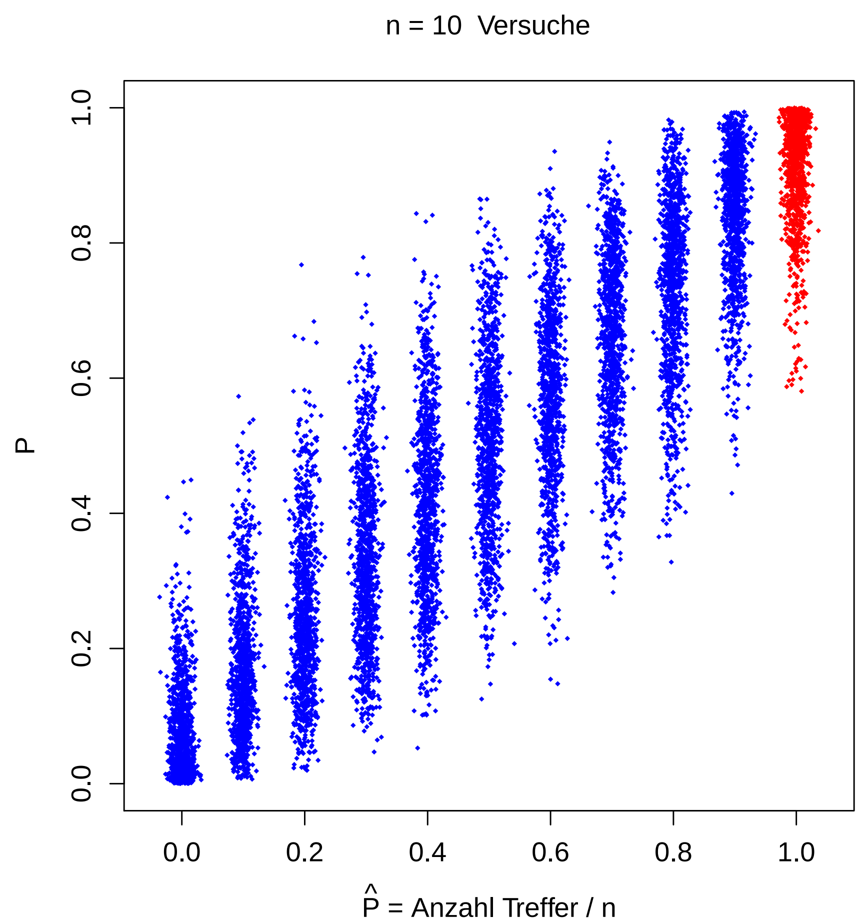
Jeder der 11000 Punkte

gibt eine gemeinsame Realisierung

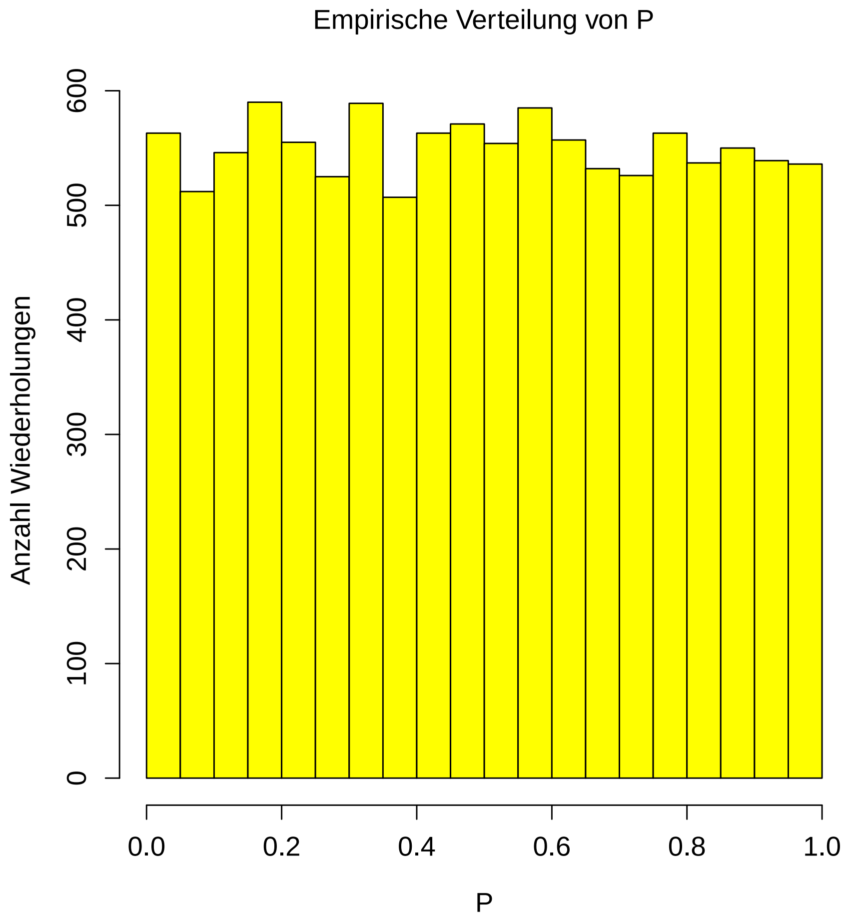
von $P$ (vertikal) und $\widehat{P} := \frac{K_{10}}{10}$ (horizontal).

Der Wertebereich von $\widehat{P}$ ist jetzt $\{0,\ 0.1, \ldots, 0.9,\ 1\}$.

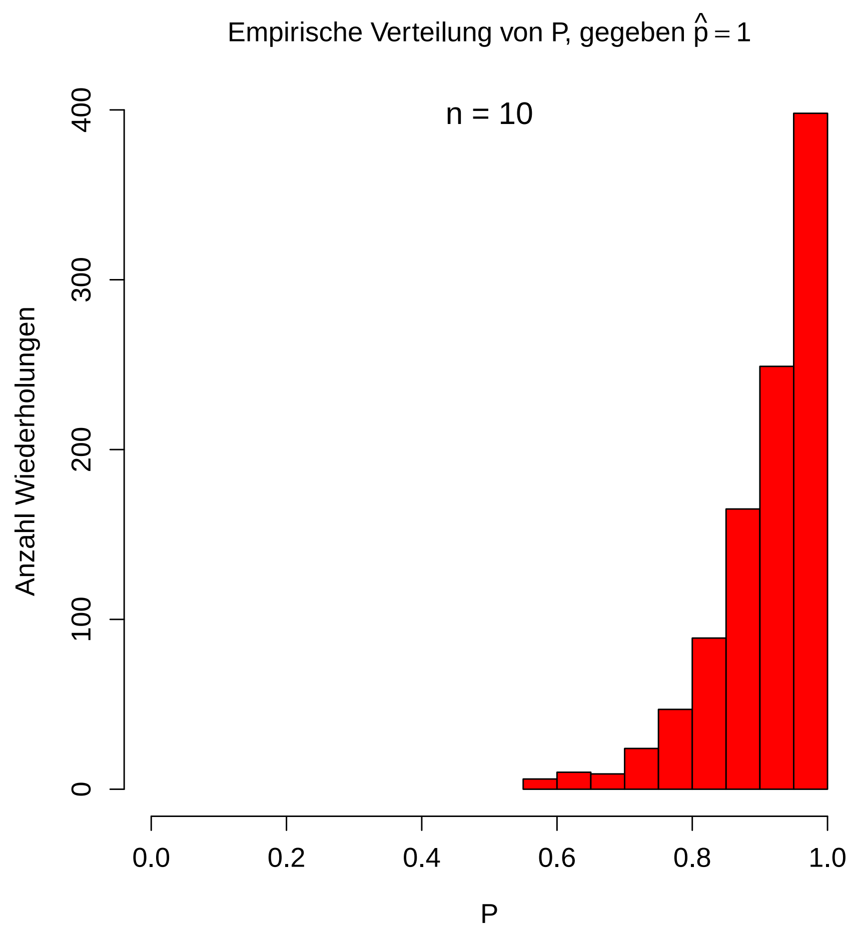
# Gemeinsame Verteilung von $(P, \frac{K_{10}}{10})$



# A-Priori Verteilung von $P$



# Bedingte Verteilung von $P$ gegeben $\{K_{10} = 10\}$



Wie kann man das elegant verstehen und “nachrechnen”?

Betrachten wir den Fall $n = 2$:

Stellen wir $(P, Z_1, Z_2)$ dar

mittlels dreier

unabhängiger, uniform auf $[0, 1]$ verteilter Zufallsvariabler

$U_0, U_1, U_2$:

$$P := U_0, \quad Z_1 := I_{\{U_1 < U_0\}}, \quad Z_2 := I_{\{U_2 < U_0\}}$$

Bis auf ein Ereignis von Wahrscheinlichkeit Null gilt:

$$\{K_2 = 2\} = \{U_0 = \max(U_0, U_1, U_2)\}.$$

Wie ist das Maximum von 3 unabhängigen

Unif([0, 1])-verteilten ZV'en verteilt?

Die Verteilungsfunktion ist

$$\mathbf{P}(\max(U_0, U_1, U_2) \le b) = b^3, \quad 0 \le b \le 1.$$

Die Dichte ist $3b^2\, db, \;\; 0 \le b \le 1$.

Das ist auch die bedingte Dichte von $P$, gegeben $\{K_2 = 2\}$.

Was $n = 2$ recht ist, soll einem allgemeinen $n$ billig sein:

Die Verteilungsfunktion von $\max(U_0, U_1, \ldots, U_n)$ ist

$$\mathbf{P}(\max(U_0, U_1, \ldots, U_n) \le b) = b^{n+1}, \quad 0 \le b \le 1.$$

Die Dichte ist $\quad (n+1)b^n\, db, \; 0 \le b \le 1.$

Das ist auch die bedingte Dichte von $P$, gegeben $\{K_n = n\}$.

Zur Beantwortung der Frage 2 für allgemeines $n$ und $k$ verwenden wir wieder die Darstellung

$$P := U_0, \qquad Z_i := I_{\{U_i < U_0\}}.$$

Erst einmal stellen wir fest:

$$\mathbf{P}(K_n = k)$$
$$= \mathbf{P}(U_0 \text{ ist das } (k+1)\text{-t kleinste der } U_0, \ldots, U_n) = \frac{1}{n+1}.$$

Die gemeinsame Verteilung von $P$ und $K_n$ ist gegeben durch

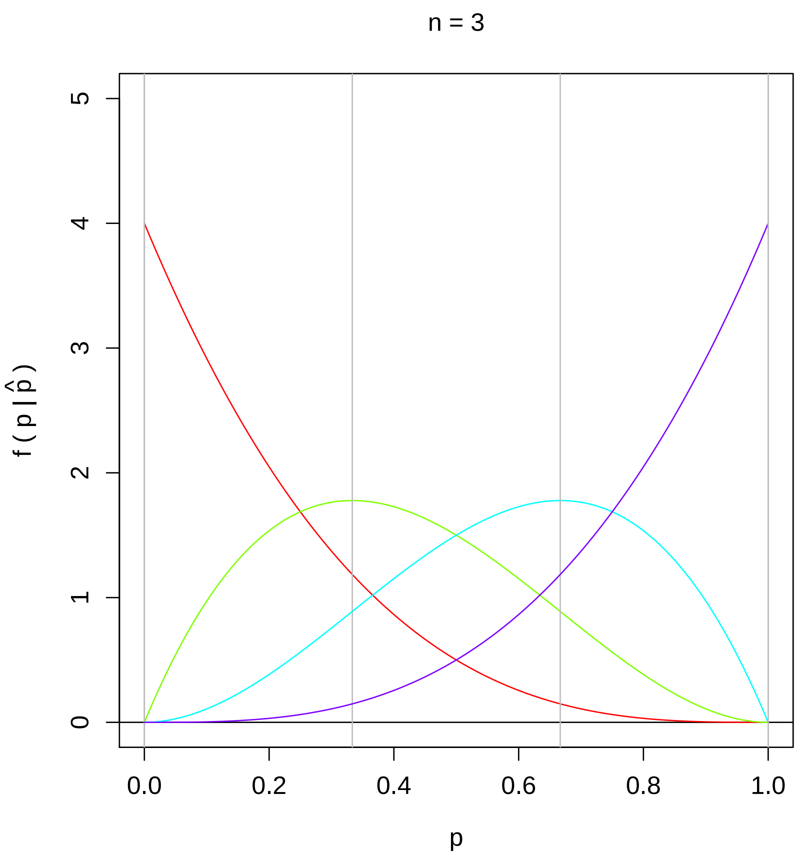
$$\mathbf{P}(P \in dp, K_n = k) = dp \binom{n}{k} p^k (1-p)^{n-k}.$$

Aus den beiden vorigen Identitäten bekommen wir die Antwort auf die Frage 2:

$$\mathbf{P}(P \in dp \,|\, K_n = k) = \frac{\mathbf{P}(P \in dp,\, K_n = k)}{\mathbf{P}(K_n = k)}$$

$$= \frac{dp \binom{n}{k} p^k (1-p)^{n-k}}{\frac{1}{n+1}} = (n+1)\binom{n}{k} p^k (1-p)^{n-k}\, dp.$$

Bedingte Dichten von $P$, gegeben $\{K_3 = k\}$, $k = 0, 1, 2, 3$:



Durch sukzessives Beobachten von immer mehr der $Z_i$

bekommt man immer mehr Information

über das à-priori uniform-verteilte $P$.

Die à posteriori Verteilung von $P$ ist immer mehr konzentriert

um einen von vornherein zufälligen Wert.

Dies wird illustriert durch die folgende Abbildung,

die an unserer Antwort

auf die eingangs gestellte Frage 2 anschließt:

# Bedingte Dichten von $P$ entlang zufälligem $(Z_1, Z_2, ...)$: