Vorlesung 9a

Bedingte Erwartung

1. Zerlegung eines Erwartungswertes nach der ersten Stufe

(Buch S. 91)

Wie in der vorigen Vorlesung betrachten wir die

gemeinsame Verteilung von zwei Zufallsvariablen X_1, X_2,

aufgebaut aus der Verteilung von X_1

und den Übergangsverteilungen:

$$\mathbf{P}(X_1 = a_1, X_2 = a_2) = \mathbf{P}(X_1 = a_1)\, \mathbf{P}_{a_1}(X_2 = a_2)$$

$$\nu(a_1, a_2) = \rho(a_1) P(a_1, a_2)$$

Auch der Erwartungswert

einer reellwertigen Zufallsvariablen $g(X_1, X_2)$

kann nach der ersten Stufe zerlegt werden.

Sei $g : S_1 \times S_2 \to \mathbb{R}$.

Wir betrachten die Zufallsvariable $h(X_1, X_2)$.

Für $a_1 \in S_1$ setzen wir

$$\mathbf{E}_{a_1}\big[g(X_1, X_2)\big] := \sum_{a_2 \in S_2} g(a_1, a_2)\mathbf{P}_{a_1}(X_2 = a_2)$$

und nennen diese Zahl den

bedingten Erwartungswert von $g(X_1, X_2)$,

gegeben $\{X_1 = a_1\}$.

Merke:

Der bedingte Erwartungswert

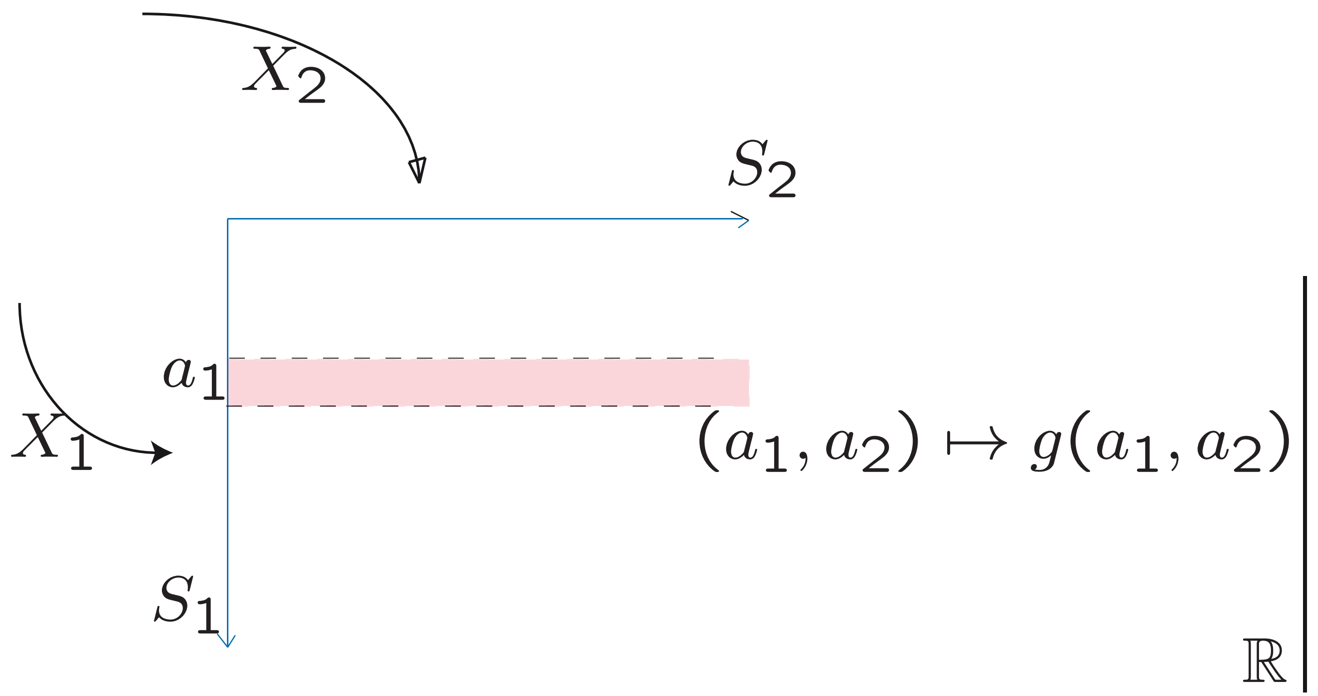
$$\mathbf{E}_{a_1}\big[g(X_1, X_2)\big]$$

wird gebildet mit der Übergangsverteilung $P(a_1, .)$,

also mit den Wahrscheinlichkeitsgewichten,

die die Zeile $P(a_1, .)$ der Matrix P bilden:

$$\mathbf{E}_{a_1}\big[g(X_1, X_2)\big] = \sum_{a_2 \in S_2} g(a_1, a_2) P(a_1, a_2)$$


X_2
S_2
a_1
X_1
$(a_1, a_2) \mapsto g(a_1, a_2)$
S_1
$\mathbb{R}$

Ähnlich wie die Verteilungsgewichte von (X_1, X_2)

lässt sich auch der Erwartungswert $\mathbf{E}[g(X_1, X_2)]$

nach den Ausgängen von X_1 zerlegen.

(Zerlegung des Erwartungswerts nach der ersten Stufe)

$$\mathbf{E}\big[g(X_1, X_2)\big]$$

$$= \sum_{a_1 \in S_1} \sum_{a_2 \in S_2} g(a_1, a_2)\, \mathbf{P}(X_1 = a_1, X_2 = a_2)$$

$$= \sum_{a_1 \in S_1} \sum_{a_2 \in S_2} g(a_1, a_2)\, \mathbf{P}(X_1 = a_1)\, \mathbf{P}_{a_1}(X_2 = a_2)$$

$$\mathbf{E}\big[g(X_1, X_2)\big]$$

$$= \sum_{a_1 \in S_1} \sum_{a_2 \in S_2} g(a_1, a_2)\, \mathbf{P}(X_1 = a_1, X_2 = a_2)$$

$$= \sum_{a_1 \in S_1} \sum_{a_2 \in S_2} g(a_1, a_2)\, \mathbf{P}(X_1 = a_1)\, \mathbf{P}_{a_1}(X_2 = a_2)$$

$$\mathbf{E}\Big[g(X_1, X_2)\Big]$$

$$= \sum_{a_1 \in S_1} \sum_{a_2 \in S_2} g(a_1, a_2)\, \mathbf{P}(X_1 = a_1, X_2 = a_2)$$

$$= \sum_{a_1 \in S_1} \sum_{a_2 \in S_2} g(a_1, a_2)\, \mathbf{P}(X_1 = a_1)\, \mathbf{P}_{a_1}(X_2 = a_2)$$

$$= \sum_{a_1 \in S_1} \mathbf{P}(X_1 = a_1) \mathbf{E}_{a_1}\Big[g(a_1, X_2)\Big]$$

$$\mathbf{E}\left[g(X_1, X_2)\right]$$

$$= \sum_{a_1 \in S_1} \sum_{a_2 \in S_2} g(a_1, a_2)\, \mathbf{P}(X_1 = a_1, X_2 = a_2)$$

$$= \sum_{a_1 \in S_1} \sum_{a_2 \in S_2} g(a_1, a_2)\, \mathbf{P}(X_1 = a_1)\, \mathbf{P}_{a_1}(X_2 = a_2)$$

$$= \sum_{a_1 \in S_1} \mathbf{P}(X_1 = a_1) \mathbf{E}_{a_1}\left[g(a_1, X_2)\right]$$

$$= \mathbf{E}\left[\mathbf{E}_{X_1}[g(X_1, X_2)]\right].$$

$$\mathbf{E}\big[g(X_1, X_2)\big]$$

$$= \sum_{a_1 \in S_1} \mathbf{P}(X_1 = a_1)\mathbf{E}_{a_1}\big[g(a_1, X_2)\big]$$

$$= \mathbf{E}\big[\mathbf{E}_{X_1}[g(X_1, X_2)]\big].$$

"Zerlegung des Erwartungswertes nach der ersten Stufe"

Merke: Der bedingte Erwartungswert

$$\mathbf{E}_{a_1}\left[g(X_1, X_2)\right]$$

ist eine Zahl.

$$\mathbf{E}_{X_1}\left[g(X_1, X_2)\right]$$

ist eine Zufallsvariable.

Wir nennen diese Zufallsvariable die

bedingte Erwartung von $g(X_1, X_2)$ gegeben X_1.

$$\mathbf{E}\big[g(X_1, X_2)\big] = \mathbf{E}\big[\mathbf{E}_{X_1}[g(X_1, X_2)]\big]$$

Ist X_2 reellwertig (also $S_2 \subset \mathbb{R}$),

dann ergibt sich als Spezialfall (mit $g(a_1, a_2) := a_2$)

$$\mathbf{E}_{a_1}[X_2] = \sum_{a_2 \in S_2} a_2 \, \mathbf{P}_{a_1}(X_2 = a_2) \;,$$

Wir haben dann die einprägsame Formel

$$\boxed{\mathbf{E}[X_2] = \mathbf{E}[\mathbf{E}_{X_1}[X_2]]}$$

(Zerlegung des Erwartungswertes von X_2 nach X_1.)

2. Ein Beispiel: Suchen in Listen.

(Buch S. 85-87)

n Namen werden in ℓ Listen einsortiert. Dadurch ergibt sich ein ℓ-Tupel $k = (k_1, \ldots, k_\ell)$ von Listenlängen (eine "Besetzung" k der Plätze $1, \ldots, \ell$)

Jeder Name steht in seiner Liste Nr. j an einer der Stellen $i = 0, \ldots, k_j - 1$.

Dieses i bezeichnet man auch als (Such-)Tiefe des Namens.

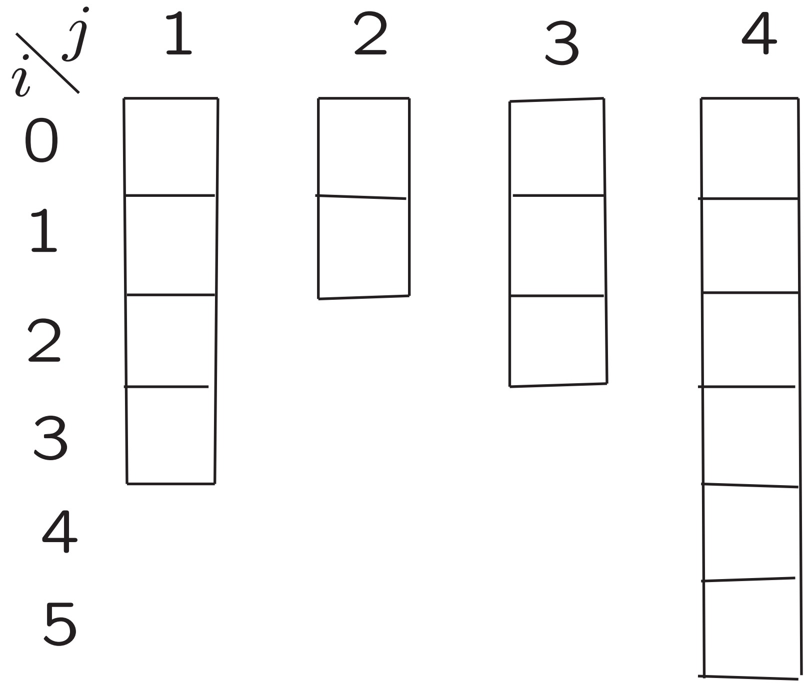
Vorstellung: Die Listennummer $j = 1, \ldots, \ell$ entspricht dem Anfangsbuchstaben des Namens.

Beispielsweise ist

für $\ell = 4$ mögliche Anfangsbuchstaben

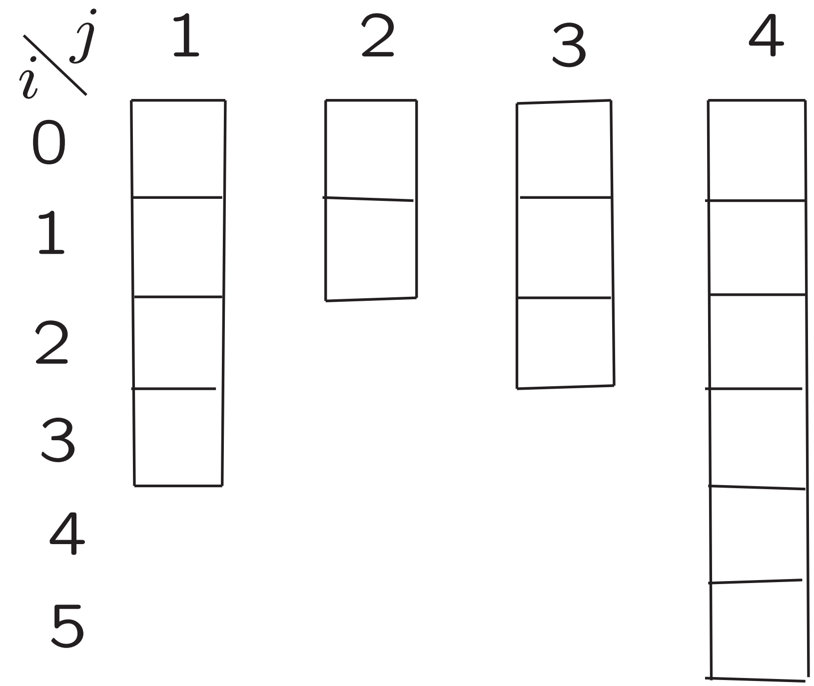
und $n = 15$ Namen

eine mögliche Besetzung:



$$n = 15, \ell = 4$$

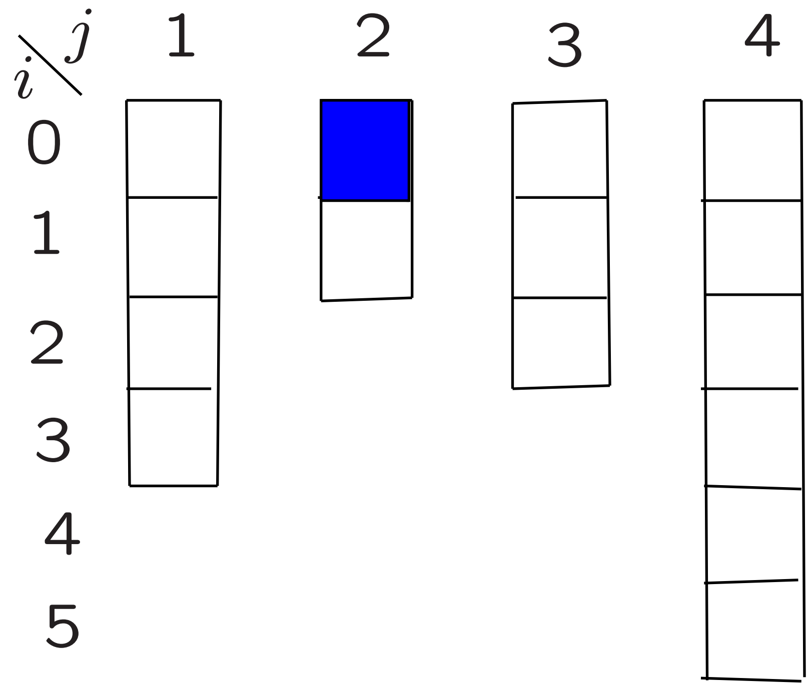
$$k = (k_1, k_2, k_3, k_4) = (4, 2, 3, 6)$$



Erste Frage: Was ist für gegebene Listenlängen $k = (k_j)$

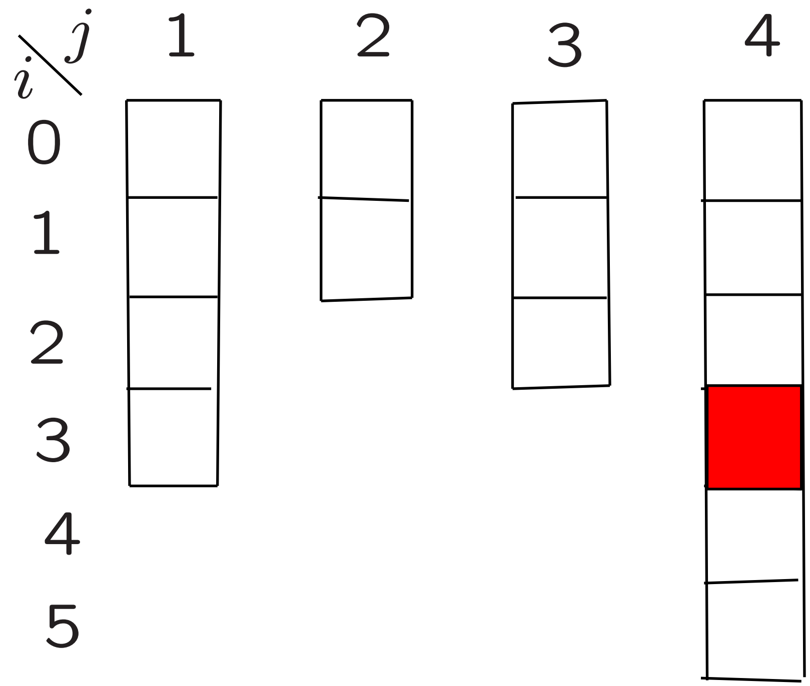
der Erwartungswert der Tiefe T

eines rein zufällig aus den n herausgegriffenen Names?



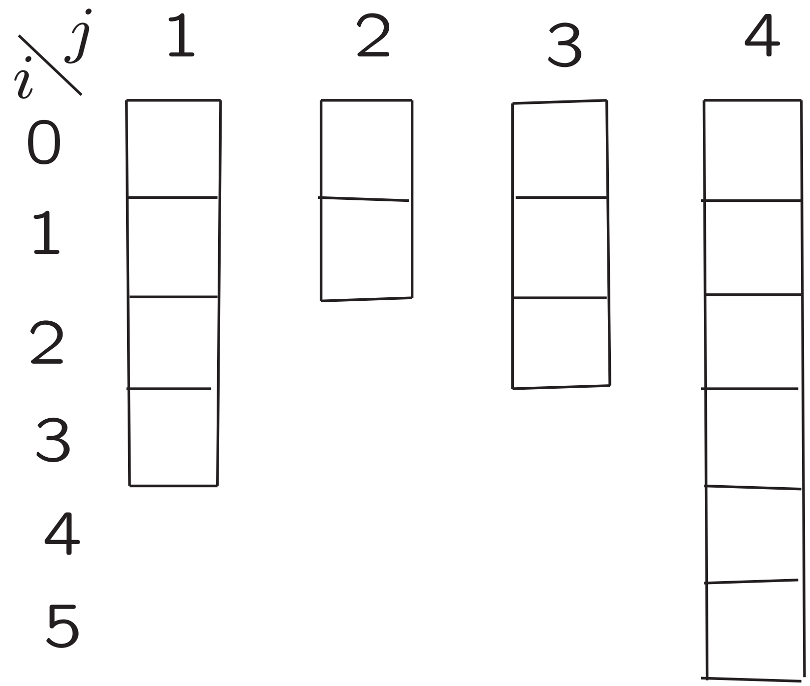
Liste $j = 2$,

Tiefe $i = 0$.



Liste $j = 4$,

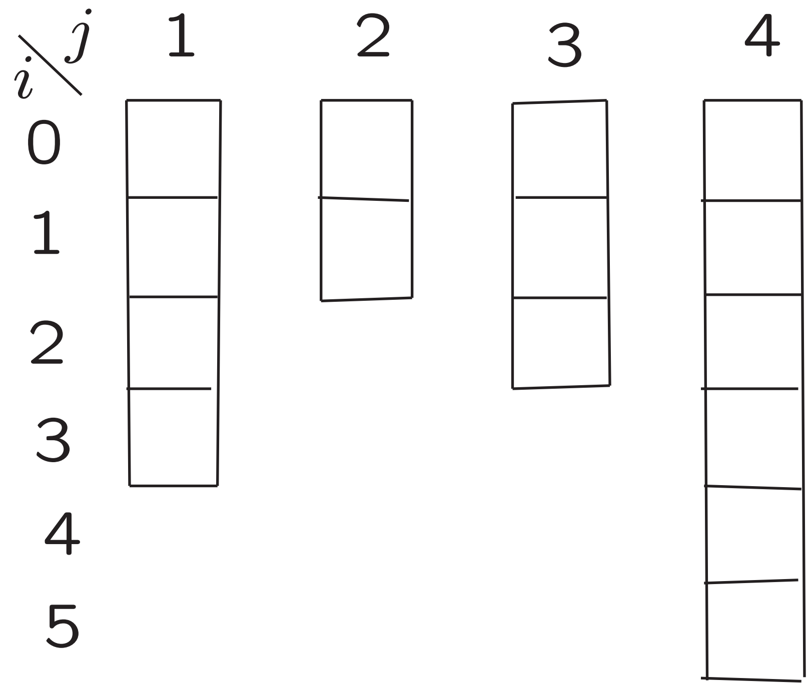
Tiefe $i = 3$.



Was ist *bei gegebenem* k

der Erwartungswert der Suchtiefe T

eines rein zufällig aus den n herausgegriffenen Names?



Die Antwort ist

$$\mathbf{E}_k[T] = \frac{1}{n}\sum_{j=1}^{\ell}\sum_{i=0}^{k_j-1} i = \frac{1}{n}\sum_{j=1}^{\ell}\frac{k_j(k_j-1)}{2}\,.$$

Wir betrachten jetzt ein

stochastisches Modell für die erste Stufe:

Annahme:

Die zufällige Besetzung $Z = (Z_1, \ldots, Z_\ell)$

kommt durch n-maliges Würfeln

mit den Gewichten $p_1, \ldots, p_\ell$ zustande.

Z ist multinomial (n, $p_1, \ldots, p_\ell$)-verteilt.

(Vorstellung: Die n Namen

sind eine Stichprobe aus einer großen Population

mit bekannten Anteilen der Anfangsbuchstaben.)

Aus den n Namen wird rein zufällig einer herausgegriffen.

Es sei T die Tiefe des Namens in seiner Liste.

Aufgabe: Berechne $\mathbf{E}[T]$.

Diese Erwartungswert beschreibt die mittlere Suchzeit

(Suchtiefe)

aller in den Listen vorhandenen Namen.

Der Erwartungswert von T, gegeben $Z = k$, war

$$\mathbf{E}_k[T] = \frac{1}{n} \sum_{j=1}^{\ell} \frac{k_j(k_j - 1)}{2} \, .$$

Mit der oben hergeleitetet Zerlegung des Erwartungswertes

$$\mathbf{E}[T] = \mathbf{E}\big[\mathbf{E}_Z[T]\big]$$

erhalten wir

$$\mathbf{E}[T] = \frac{1}{n} \sum_{j=1}^{\ell} \mathbf{E}\Big[\frac{Z_j(Z_j - 1)}{2}\Big] \, .$$

$$\mathbf{E}[T] = \frac{1}{n} \sum_{j=1}^{\ell} \mathbf{E}\left[\frac{Z_j(Z_j - 1)}{2}\right]$$

Nach Annahme ist Z_j Binomial(n, p_j)-verteilt.

Mit der Formel $\mathbf{Var}[Z] = \mathbf{E}[Z^2] - (\mathbf{E}[Z])^2$ folgt

$$\mathbf{E}[Z_j(Z_j - 1)] = \mathbf{Var}[Z_j] + \mathbf{E}[Z_j]^2 - \mathbf{E}[Z_j]$$

$$= np_j(1 - p_j) + (np_j)^2 - np_j = p_j^2\, n(n-1),$$

$$\mathbf{E}[T] = \frac{n-1}{2}(p_1^2 + \cdots + p_\ell^2)\ .$$

$$\mathbf{E}[T] = \frac{n-1}{2}(p_1^2 + \cdots + p_\ell^2) \ .$$

Im Fall uniformer Gewichte

$$p_1 = \cdots = p_\ell = 1/\ell$$

ergibt sich

$$\mathbf{E}[T] = \frac{n-1}{2\ell} \ .$$

Eine Modifikation des vorigen Beispiels:

Sei Z wieder multinomial $(n, p_1, \dots, p_\ell)$-verteilt,
J sei unabhängig von Z, mit $\mathbf{P}(J = j) = p_j, \quad j = 1, \dots, \ell$.

Berechne den Erwartungswert von $G := Z_J$.

(Man kann dabei denken an den Erwartungswert der Suchzeit nach einem in den Listen nicht vorhandenen Namen.)

Wir fassen Z als erste Stufe auf, und Z_J als zweite, und . . .

…zerlegen $\mathbf{E}[Z_J]$ nach den Ausgängen von Z:

$$\mathbf{E}[G] = \mathbf{E}[Z_J] = \sum_k \mathbf{P}(Z = k)\, \mathbf{E}_k[Z_J]$$

$$= \sum_k \mathbf{P}(Z = k)\, \mathbf{E}_k[k_J] = \sum_k \mathbf{P}(Z = k) \sum_{j=1}^{\ell} k_j\, p_j$$

$$= \sum_{j=1}^{\ell} p_j \sum_k k_j\, \mathbf{P}(Z = k)$$

$$= \sum_{j=1}^{\ell} p_j\, \mathbf{E}Z_j = \sum_{j=1}^{\ell} p_j\, np_j = n \sum_{j=1}^{\ell} p_j^2 .$$

Im Fall uniformer Gewichte

$$p_1 = \cdots = p_\ell = 1/\ell$$

ergibt sich

$$\mathbf{E}[G] = \frac{n}{\ell}.$$

Im Vergleich dazu war (siehe voriges Besipiel)

die erwartete Suchtiefe eines rein zufällig aus den n

herausgegriffenen Namens

$$\mathbf{E}[T] = \frac{n-1}{2\ell}\,.$$