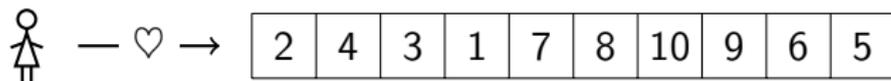


Das Heiratsproblem

Optimal Stopping

Das Modell

- ▶ Im Laufe unseres Lebens lernen wir mehrere potenzielle Partner kennen.
- ▶ Wir möchten den oder die „optimale(n)“ Partner(in) finden. Dazu und ordnen wir jedem Kandidaten einen Rang zu.
- ▶ Das Problem: Wir können nicht in die Zukunft sehen und wir können später auch nicht zu einem verlassenen Partner zurück.



Eine präzise Formulierung der Problemstellung

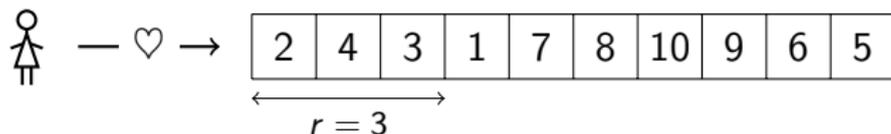
Die folgenden Formulierungen sind willkürlich aus der Sicht einer Frau gewählt.

- ▶ Es gibt n Kandidaten und die Zahl n ist im voraus bekannt.
- ▶ Es wird genau ein Partner gewählt.
- ▶ Die Bewerber können, wenn sie alle zusammen gesehen werden, in aufsteigender Rangfolge eindeutig vom schlechtesten zum besten geordnet werden.
- ▶ Die Bewerber werden alle nacheinander in zufälliger Reihenfolge begutachtet, wobei jede Reihenfolge gleich wahrscheinlich ist.
- ▶ Nachdem man sich über einen potenzieller Partner ein Bild gemacht hat, trennt man sich oder man entscheidet sich für die Heirat und die Entscheidung ist unwiderruflich.

- ▶ Die Entscheidung, einen Bewerber anzunehmen oder abzulehnen, kann nur auf den relativen Rängen der bisher kennen gelernten Bewerbern beruhen.
- ▶ Das Ziel der allgemeinen Lösung ist es, mit der höchsten Wahrscheinlichkeit den besten Partner der gesamten Gruppe auszuwählen.

Die von uns untersuchte Strategie

- ▶ „Teste“ die ersten r potenziellen Partner, ordne ihnen einen eindeutigen Rang zu und verlasse sie danach.
- ▶ Wähle von den verbleibenden Bewerbern den ersten, der besser ist als jeder der ersten r . Wenn es keinen besseren gibt, wähle den letzten.

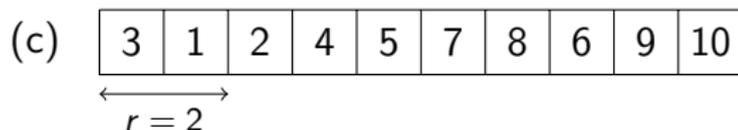
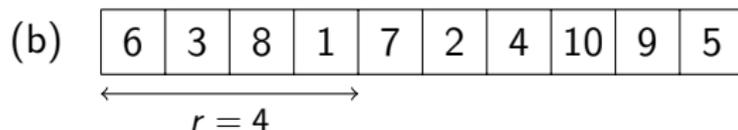
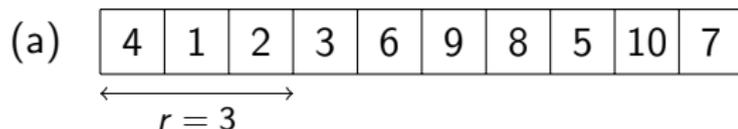


Die zentrale Frage für den Rest dieses Dokuments

Wie gross muss r gewählt werden, damit die Chance auf den *besten* Partner (im obigen Beispiel ist es der mit dem Rang 10) maximal wird?

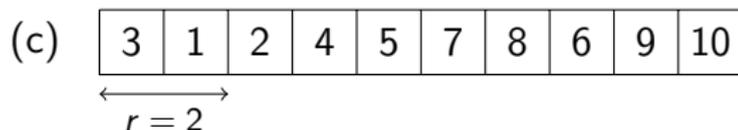
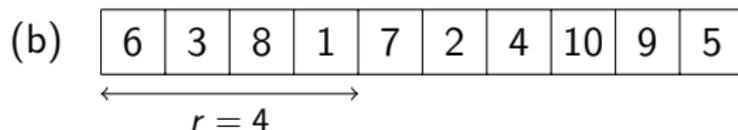
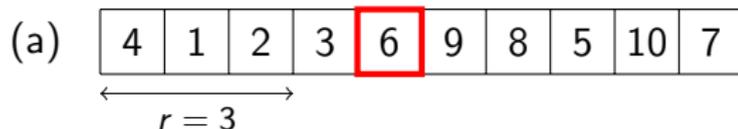
Aufgabe 1

Bestimme den Rang des Partners, den eine Partnersuchende nach der oben beschriebenen Strategie wählen muss.



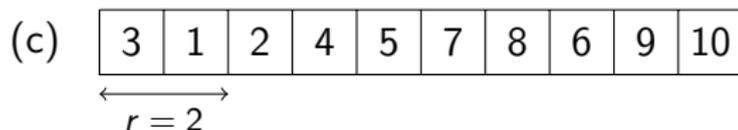
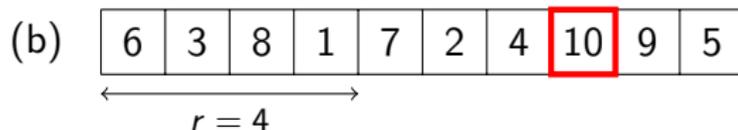
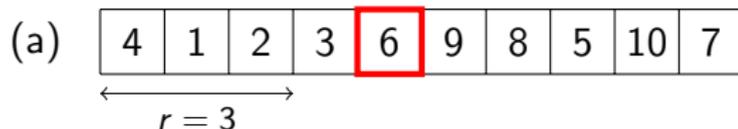
Aufgabe 1

Bestimme den Rang des Partners, den eine Partnersuchende nach der oben beschriebenen Strategie wählen muss.



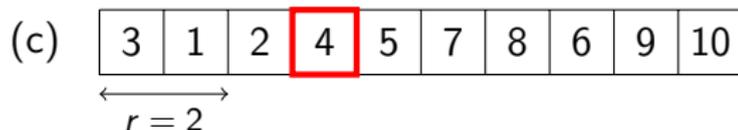
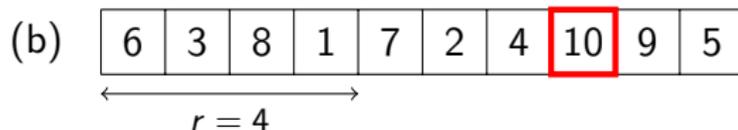
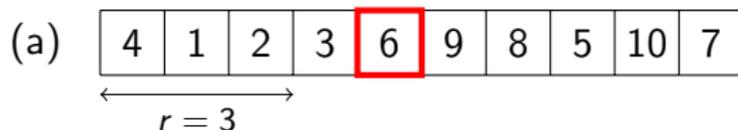
Aufgabe 1

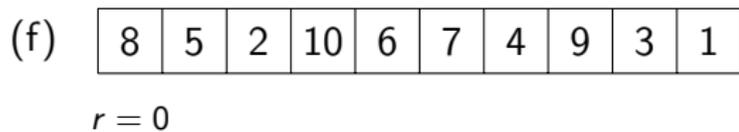
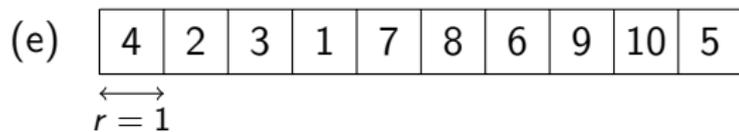
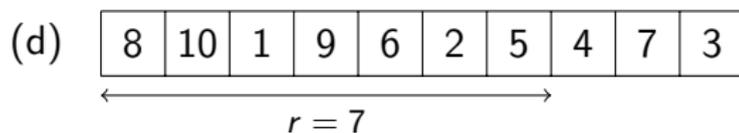
Bestimme den Rang des Partners, den eine Partnersuchende nach der oben beschriebenen Strategie wählen muss.

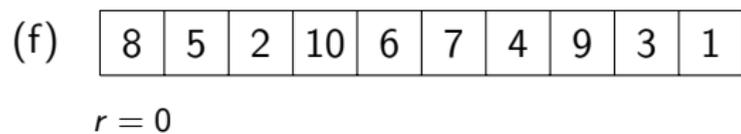
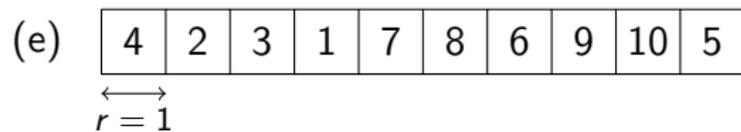
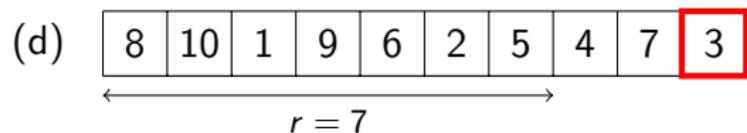


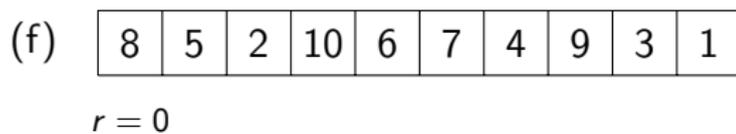
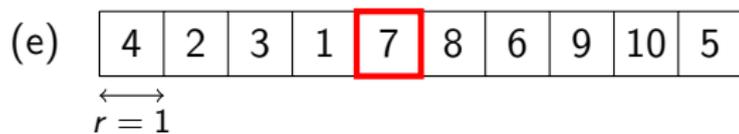
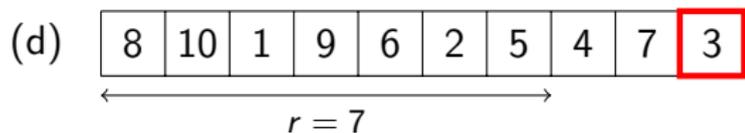
Aufgabe 1

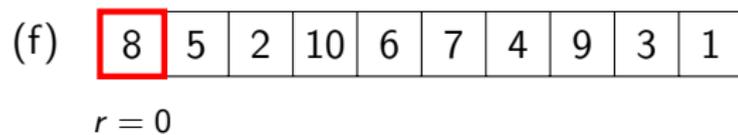
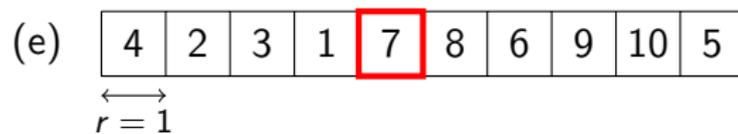
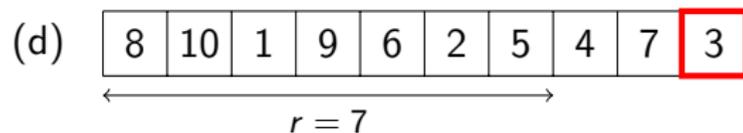
Bestimme den Rang des Partners, den eine Partnersuchende nach der oben beschriebenen Strategie wählen muss.











Aufgabe 2

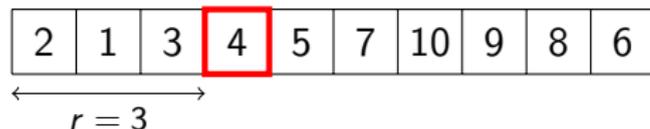
Beschreibe, warum die folgende Situation für die partnersuchende Person ungünstig ist.

2	1	3	4	5	7	10	9	8	6
---	---	---	---	---	---	----	---	---	---

\longleftarrow
 $r = 3$

Aufgabe 2

Beschreibe, warum die folgende Situation für die partnersuchende Person ungünstig ist.



Man lernt in der Testphase nur Kandidaten mit einem schlechten „Score“ kennen und gerät dann gleich danach an jemanden, der nur ein bisschen besser ist. Diese(r) muss dann gewählt werden, wenn man den Algorithmus befolgt.

Aufgabe 3

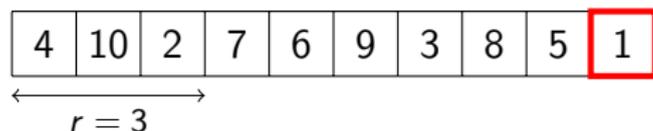
Beschreibe, warum die folgende Situation für die partnersuchende Person ungünstig ist.

4	10	2	7	6	9	3	8	5	1
---	----	---	---	---	---	---	---	---	---

\longleftarrow $r = 3$ \longrightarrow

Aufgabe 3

Beschreibe, warum die folgende Situation für die partnersuchende Person ungünstig ist.



Man lernt bereits in der Testphase den optimalen Partner kennen, muss ihn aber wieder verlassen, wenn man den Algorithmus befolgt. In diese Fall kann man keinen besseren finden und landet beim letzten in der Liste.

Aufgabe 4

Die folgende Tabelle enthält *alle* möglichen Reihenfolgen von 4 Kandidaten (Rängen).

- (a) Warum gibt es 24 Reihenfolgen?
- (b) Vervollständige die Tabelle, indem du jeweils den Rang des Kandidaten bestimmst, den wir mit der oben beschriebenen Strategie, nach $r = 0, 1, 2, 3$ begutachteten Kandidaten wählen müssen. Zwei Zeilen sind bereits ausgefüllt.

Aufgabe 4

Die folgende Tabelle enthält *alle* möglichen Reihenfolgen von 4 Kandidaten (Rängen).

- (a) Warum gibt es 24 Reihenfolgen? $4 \cdot 3 \cdot 2 \cdot 1 = 24 = 4!$
- (b) Vervollständige die Tabelle, indem du jeweils den Rang des Kandidaten bestimmst, den wir mit der oben beschriebenen Strategie, nach $r = 0, 1, 2, 3$ begutachteten Kandidaten wählen müssen. Zwei Zeilen sind bereits ausgefüllt.

(c) Für welches r erhält man am häufigsten den Rang 4?

Reihenfolge	bester nach $r =$				Reihenfolge	bester nach $r =$			
	0	1	2	3		0	1	2	3
1, 2, 3, 4	1	2	3	4	3, 1, 2, 4	3	4	4	4
1, 2, 4, 3					3, 1, 4, 2				
1, 3, 2, 4					3, 2, 1, 4				
1, 3, 4, 2					3, 2, 4, 1				
1, 4, 2, 3					3, 4, 1, 2				
1, 4, 3, 2					3, 4, 2, 1				
2, 1, 3, 4					4, 1, 2, 3				
2, 1, 4, 3					4, 1, 3, 2				
2, 3, 1, 4					4, 2, 1, 3				
2, 3, 4, 1					4, 2, 3, 1				
2, 4, 1, 3					4, 3, 1, 2				
2, 4, 3, 1					4, 3, 2, 1				
	$r = 0$	$r = 1$	$r = 2$	$r = 3$					
Anzahl Rang 4									

(c) Für welches r erhält man am häufigsten den Rang 4?

Reihenfolge	bester nach $r =$				Reihenfolge	bester nach $r =$			
	0	1	2	3		0	1	2	3
1, 2, 3, 4	1	2	3	4	3, 1, 2, 4	3	4	4	4
1, 2, 4, 3	1				3, 1, 4, 2	3			
1, 3, 2, 4	1				3, 2, 1, 4	3			
1, 3, 4, 2	1				3, 2, 4, 1	3			
1, 4, 2, 3	1				3, 4, 1, 2	3			
1, 4, 3, 2	1				3, 4, 2, 1	3			
2, 1, 3, 4	2				4, 1, 2, 3	4			
2, 1, 4, 3	2				4, 1, 3, 2	4			
2, 3, 1, 4	2				4, 2, 1, 3	4			
2, 3, 4, 1	2				4, 2, 3, 1	4			
2, 4, 1, 3	2				4, 3, 1, 2	4			
2, 4, 3, 1	2				4, 3, 2, 1	4			
	$r = 0$	$r = 1$	$r = 2$	$r = 3$					
Anzahl Rang 4									

(c) Für welches r erhält man am häufigsten den Rang 4?

Reihenfolge	bester nach $r =$				Reihenfolge	bester nach $r =$			
	0	1	2	3		0	1	2	3
1, 2, 3, 4	1	2	3	4	3, 1, 2, 4	3	4	4	4
1, 2, 4, 3	1	2			3, 1, 4, 2	3	4		
1, 3, 2, 4	1	3			3, 2, 1, 4	3	4		
1, 3, 4, 2	1	3			3, 2, 4, 1	3	4		
1, 4, 2, 3	1	4			3, 4, 1, 2	3	4		
1, 4, 3, 2	1	4			3, 4, 2, 1	3	4		
2, 1, 3, 4	2	3			4, 1, 2, 3	4	3		
2, 1, 4, 3	2	4			4, 1, 3, 2	4	2		
2, 3, 1, 4	2	3			4, 2, 1, 3	4	3		
2, 3, 4, 1	2	3			4, 2, 3, 1	4	1		
2, 4, 1, 3	2	4			4, 3, 1, 2	4	2		
2, 4, 3, 1	2	4			4, 3, 2, 1	4	1		
	$r = 0$	$r = 1$	$r = 2$	$r = 3$					
Anzahl Rang 4									

(c) Für welches r erhält man am häufigsten den Rang 4?

Reihenfolge	bester nach $r =$				Reihenfolge	bester nach $r =$			
	0	1	2	3		0	1	2	3
1, 2, 3, 4	1	2	3	4	3, 1, 2, 4	3	4	4	4
1, 2, 4, 3	1	2	4		3, 1, 4, 2	3	4	4	
1, 3, 2, 4	1	3	4		3, 2, 1, 4	3	4	4	
1, 3, 4, 2	1	3	4		3, 2, 4, 1	3	4	4	
1, 4, 2, 3	1	4	3		3, 4, 1, 2	3	4	2	
1, 4, 3, 2	1	4	2		3, 4, 2, 1	3	4	1	
2, 1, 3, 4	2	3	3		4, 1, 2, 3	4	3	3	
2, 1, 4, 3	2	4	4		4, 1, 3, 2	4	2	2	
2, 3, 1, 4	2	3	4		4, 2, 1, 3	4	3	3	
2, 3, 4, 1	2	3	4		4, 2, 3, 1	4	1	1	
2, 4, 1, 3	2	4	3		4, 3, 1, 2	4	2	2	
2, 4, 3, 1	2	4	1		4, 3, 2, 1	4	1	1	
	$r = 0$	$r = 1$	$r = 2$	$r = 3$					
Anzahl Rang 4									

(c) Für welches r erhält man am häufigsten den Rang 4?

Reihenfolge	bester nach $r =$				Reihenfolge	bester nach $r =$			
	0	1	2	3		0	1	2	3
1, 2, 3, 4	1	2	3	4	3, 1, 2, 4	3	4	4	4
1, 2, 4, 3	1	2	4	3	3, 1, 4, 2	3	4	4	2
1, 3, 2, 4	1	3	4	4	3, 2, 1, 4	3	4	4	4
1, 3, 4, 2	1	3	4	2	3, 2, 4, 1	3	4	4	1
1, 4, 2, 3	1	4	3	3	3, 4, 1, 2	3	4	2	2
1, 4, 3, 2	1	4	2	2	3, 4, 2, 1	3	4	1	1
2, 1, 3, 4	2	3	3	4	4, 1, 2, 3	4	3	3	3
2, 1, 4, 3	2	4	4	3	4, 1, 3, 2	4	2	2	2
2, 3, 1, 4	2	3	4	4	4, 2, 1, 3	4	3	3	3
2, 3, 4, 1	2	3	4	1	4, 2, 3, 1	4	1	1	1
2, 4, 1, 3	2	4	3	3	4, 3, 1, 2	4	2	2	2
2, 4, 3, 1	2	4	1	1	4, 3, 2, 1	4	1	1	1
	$r = 0$	$r = 1$	$r = 2$	$r = 3$					
Anzahl Rang 4									

(c) Für welches r erhält man am häufigsten den Rang 4?

Reihenfolge	bester nach $r =$				Reihenfolge	bester nach $r =$			
	0	1	2	3		0	1	2	3
1, 2, 3, 4	1	2	3	4	3, 1, 2, 4	3	4	4	4
1, 2, 4, 3	1	2	4	3	3, 1, 4, 2	3	4	4	2
1, 3, 2, 4	1	3	4	4	3, 2, 1, 4	3	4	4	4
1, 3, 4, 2	1	3	4	2	3, 2, 4, 1	3	4	4	1
1, 4, 2, 3	1	4	3	3	3, 4, 1, 2	3	4	2	2
1, 4, 3, 2	1	4	2	2	3, 4, 2, 1	3	4	1	1
2, 1, 3, 4	2	3	3	4	4, 1, 2, 3	4	3	3	3
2, 1, 4, 3	2	4	4	3	4, 1, 3, 2	4	2	2	2
2, 3, 1, 4	2	3	4	4	4, 2, 1, 3	4	3	3	3
2, 3, 4, 1	2	3	4	1	4, 2, 3, 1	4	1	1	1
2, 4, 1, 3	2	4	3	3	4, 3, 1, 2	4	2	2	2
2, 4, 3, 1	2	4	1	1	4, 3, 2, 1	4	1	1	1
	$r = 0$	$r = 1$	$r = 2$	$r = 3$					
Anzahl Rang 4	6								

(c) Für welches r erhält man am häufigsten den Rang 4?

Reihenfolge	bester nach $r =$				Reihenfolge	bester nach $r =$			
	0	1	2	3		0	1	2	3
1, 2, 3, 4	1	2	3	4	3, 1, 2, 4	3	4	4	4
1, 2, 4, 3	1	2	4	3	3, 1, 4, 2	3	4	4	2
1, 3, 2, 4	1	3	4	4	3, 2, 1, 4	3	4	4	4
1, 3, 4, 2	1	3	4	2	3, 2, 4, 1	3	4	4	1
1, 4, 2, 3	1	4	3	3	3, 4, 1, 2	3	4	2	2
1, 4, 3, 2	1	4	2	2	3, 4, 2, 1	3	4	1	1
2, 1, 3, 4	2	3	3	4	4, 1, 2, 3	4	3	3	3
2, 1, 4, 3	2	4	4	3	4, 1, 3, 2	4	2	2	2
2, 3, 1, 4	2	3	4	4	4, 2, 1, 3	4	3	3	3
2, 3, 4, 1	2	3	4	1	4, 2, 3, 1	4	1	1	1
2, 4, 1, 3	2	4	3	3	4, 3, 1, 2	4	2	2	2
2, 4, 3, 1	2	4	1	1	4, 3, 2, 1	4	1	1	1
	$r = 0$	$r = 1$	$r = 2$	$r = 3$					
Anzahl Rang 4	6	11							

(c) Für welches r erhält man am häufigsten den Rang 4?

Reihenfolge	bester nach $r =$				Reihenfolge	bester nach $r =$			
	0	1	2	3		0	1	2	3
1, 2, 3, 4	1	2	3	4	3, 1, 2, 4	3	4	4	4
1, 2, 4, 3	1	2	4	3	3, 1, 4, 2	3	4	4	2
1, 3, 2, 4	1	3	4	4	3, 2, 1, 4	3	4	4	4
1, 3, 4, 2	1	3	4	2	3, 2, 4, 1	3	4	4	1
1, 4, 2, 3	1	4	3	3	3, 4, 1, 2	3	4	2	2
1, 4, 3, 2	1	4	2	2	3, 4, 2, 1	3	4	1	1
2, 1, 3, 4	2	3	3	4	4, 1, 2, 3	4	3	3	3
2, 1, 4, 3	2	4	4	3	4, 1, 3, 2	4	2	2	2
2, 3, 1, 4	2	3	4	4	4, 2, 1, 3	4	3	3	3
2, 3, 4, 1	2	3	4	1	4, 2, 3, 1	4	1	1	1
2, 4, 1, 3	2	4	3	3	4, 3, 1, 2	4	2	2	2
2, 4, 3, 1	2	4	1	1	4, 3, 2, 1	4	1	1	1
	$r = 0$	$r = 1$	$r = 2$	$r = 3$					
Anzahl Rang 4	6	11	10						

(c) Für welches r erhält man am häufigsten den Rang 4?

Reihenfolge	bester nach $r =$				Reihenfolge	bester nach $r =$			
	0	1	2	3		0	1	2	3
1, 2, 3, 4	1	2	3	4	3, 1, 2, 4	3	4	4	4
1, 2, 4, 3	1	2	4	3	3, 1, 4, 2	3	4	4	2
1, 3, 2, 4	1	3	4	4	3, 2, 1, 4	3	4	4	4
1, 3, 4, 2	1	3	4	2	3, 2, 4, 1	3	4	4	1
1, 4, 2, 3	1	4	3	3	3, 4, 1, 2	3	4	2	2
1, 4, 3, 2	1	4	2	2	3, 4, 2, 1	3	4	1	1
2, 1, 3, 4	2	3	3	4	4, 1, 2, 3	4	3	3	3
2, 1, 4, 3	2	4	4	3	4, 1, 3, 2	4	2	2	2
2, 3, 1, 4	2	3	4	4	4, 2, 1, 3	4	3	3	3
2, 3, 4, 1	2	3	4	1	4, 2, 3, 1	4	1	1	1
2, 4, 1, 3	2	4	3	3	4, 3, 1, 2	4	2	2	2
2, 4, 3, 1	2	4	1	1	4, 3, 2, 1	4	1	1	1
	$r = 0$	$r = 1$	$r = 2$	$r = 3$					
Anzahl Rang 4	6	11	10	6					

Aufgabe 5 (Python)

```
L = [3, 1, 5, 7, 4, 8, 6, 10, 2, 9]
```

```
max(L[:3])?
```

Aufgabe 5 (Python)

```
L = [3, 1, 5, 7, 4, 8, 6, 10, 2, 9]
```

```
max(L[:3])?    7
```

Das Simulationsprogramm

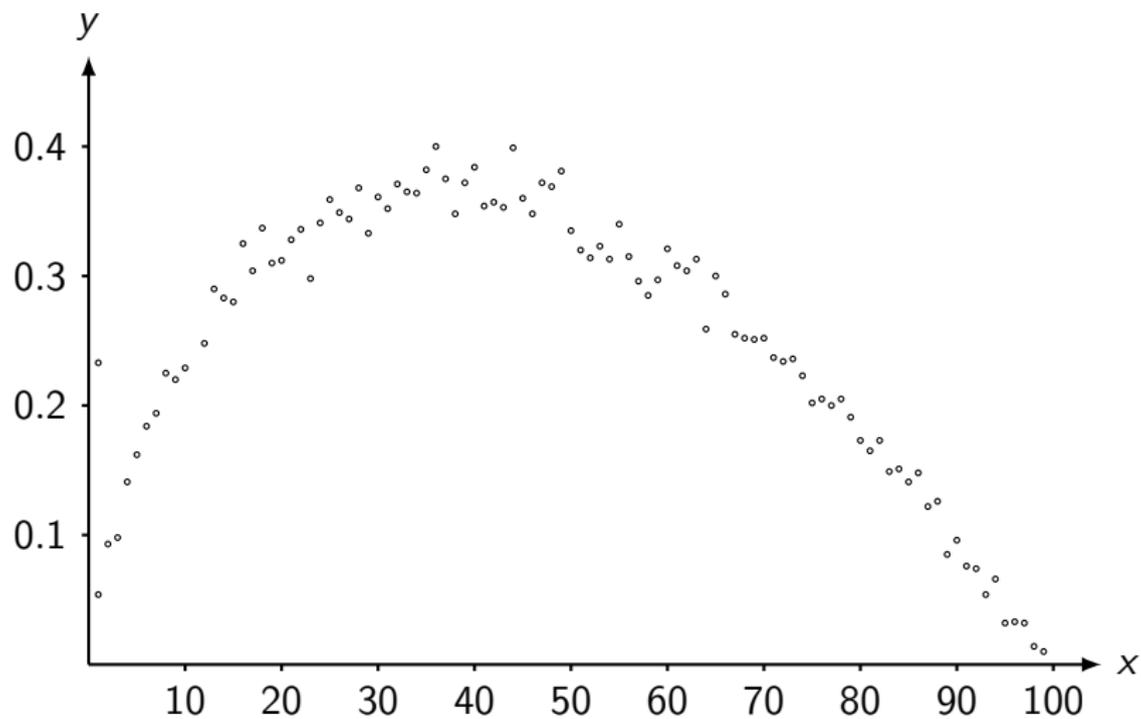
```
1 from random import sample
2
3 def erzeuge_kandidaten(n):
4     '''Gibt [1, 2, 3, ..., n] zufällig gemischt zurück.'''
5     return sample(range(1, n+1), n)
6
7 def bester_nach_test(liste, r):
8     '''Gibt besten Kandidaten bei r Testkandidaten
9         zurück.'''
10    test_max = max(liste[:r])
11    for wert in liste[r:]:
12        if wert > test_max: # ist jemand besser?
13            return wert     # wenn ja, gibt seinen Wert
14                            # zurück
15    return liste[-1]       # ... sonst den Wert des
16                            # letzten
```

Das Simulationsprogramm

```
15 def simulation(n, r, w):
16     '''Gibt Anteil der optimalen Wahl bei w
17         Wiederholungen zurück.'''
18     erfolge = 0
19     best = n
20     for i in range(0, w):
21         liste = erzeuge_kandidaten(n)
22         wahl = bester_nach_test(liste, r)
23         if wahl == best:
24             erfolge += 1
25     return erfolge/w # relative Häufigkeit
```

Das Simulationsprogramm

```
26 # Wiederhole die Simulation w-Mal für jedes r und schreibe
27 # r und die zugehörige Erfolgsquote in eine CSV-Datei.
28 dd = open('partnerwahl.csv', mode='w')
29 dd.write('r;q\n')
30 n, w = 100, 5000 # Anzahl Kandidaten & Wiederholungen
31 for r in range(1, n):
32     erfolgsquote = simulation(n, r, w)
33     print(r, erfolgsquote)
34     dd.write('{0};{1}\n'.format(r, erfolgsquote))
35 dd.close()
```



Aufgabe 6

Führe das obige Simulationsprogramm mit $n = 100$ und $w = 5000$ durch. Öffne die so erzeugte Datei `heiratsproblem.csv` in einem Tabellenkalulationsprogramm und stelle die Daten als Liniendiagramm dar (Werte der erste Kolonne \rightarrow Achsenbeschriftung).

Für welche Anteil Kandidaten (r/n) sind die Chancen, den besten Partner zu finden, maximal? Vergleiche mit:
<https://de.wikipedia.org/wiki/Sekretärinnenproblem>

Aufgabe 6

Führe das obige Simulationsprogramm mit $n = 100$ und $w = 5000$ durch. Öffne die so erzeugte Datei `heiratsproblem.csv` in einem Tabellenkalkulationsprogramm und stelle die Daten als Liniendiagramm dar (Werte der erste Kolonne \rightarrow Achsenbeschriftung).

Für welche Anteil Kandidaten (r/n) sind die Chancen, den besten Partner zu finden, maximal? Vergleiche mit:

<https://de.wikipedia.org/wiki/Sekretärinnenproblem>

Die obige Simulation lässt vermuten, dass man am häufigsten die beste Wahl erhält, wenn der Anteil r der zu testenden Kandidaten zwischen 0.36 und 0.37 liegt. Der Wert von $q = r/n$ strebt gegen $1/e \approx 0.367879\dots$, wenn man die Anzahl der Kandidaten und die Anzahl der Versuche w gegen ∞ streben lässt?