

Ue Statistik u. Wahrscheinlichkeitsth. f. Inf. 107.251 W 2005 6 http://www.statistik.tuwien.ac.at/RISueb/	Di 12:00 – 18:00
	2.Blatt
Werner GURKER Tel.: 58801-10724 Spr.: Di u. Do 11-12 e-mail: W.Gurker@tuwien.ac.at	18. Okt. 2005

Die folgenden graphischen Aufgaben können sowohl „mit der Hand“ als auch mit Hilfe von (Statistik-) Software (R, EXCEL, ...) gelöst werden.

1. Klassifizieren Sie die Merkmale dieses Übungsblatts nach den in der Vorlesung besprochenen Gesichtspunkten (qualitativ/quantitativ; diskret/stetig; ...) und bestimmen Sie das jeweilige Skalenniveau (Nominal-, Ordinal-, Intervall-, Verhältnisskala). Geben Sie auch passende Merkmalräume an.

Hinweis: *Intervallskalen* haben keinen absolut festgelegten Nullpunkt, er dient nur zur Definition der Skala; Differenzen sind sinnvoll interpretierbar (Bsp.: Temperaturskala in °C). *Verhältnisskalen* sind Intervallskalen mit absolutem Nullpunkt und nur positiven Meßwerten (Bsp.: Temperaturskala in K(elvin)).

2. Die folgende Tabelle zeigt die Benotung der UE „Statistik u. Wahrscheinlichkeitsth. f. Inf.“ vom WS 2004/5 (nur die positiven Noten), aufgeschlüsselt nach den UE-Leitern:

	S1	U2	B3	G4
L1	21	9	3	6
L2	24	30	20	0
L3	7	16	23	2
L4	8	24	25	5
L5	4	12	20	11

- (a) Zeichnen Sie für die Gesamtbenotung ein Balken- und ein Kreisdiagramm.
- (b) Zeichnen Sie für die Benotungen der UE-Leiter Balkendiagramme. Achten Sie bei der Darstellung darauf, daß die Diagramme miteinander vergleichbar sind und kommentieren Sie das Ergebnis.

Die Beispiele 3 – 6 beziehen sich auf den Datensatz **normtemp.dat** (vgl. Ü-Homepage), der Angaben zu Körpertemperatur (**temperature** [°F]; $x^{\circ}\text{F} = 5(x - 32)/9^{\circ}\text{C}$) und Ruhepuls (**hr** [Schläge pro Minute]) von 130 (gesunden) Personen beiderlei Geschlechts (**gender**) enthält. Studentinnen betrachten den Teildatensatz für **gender=W** und Studenten den für **gender=M**.

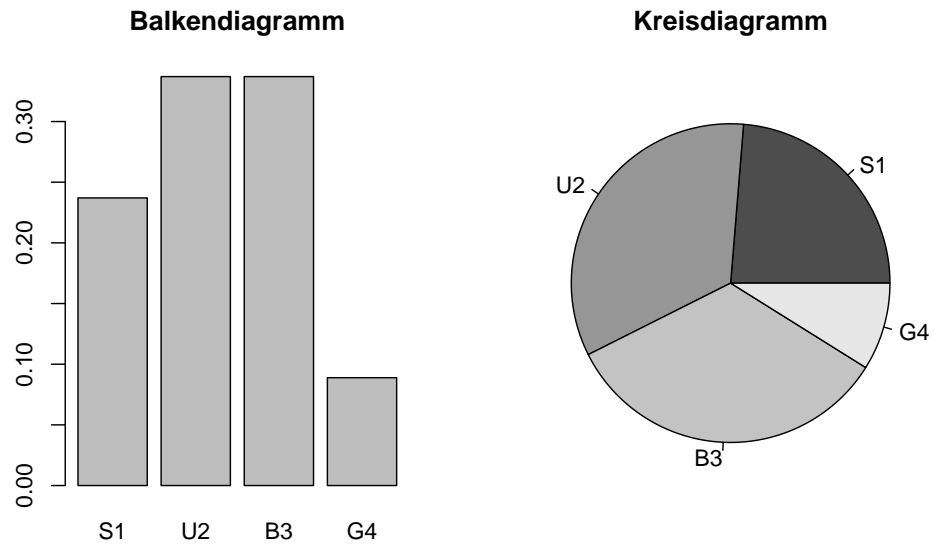
3. Zeichnen Sie für das Merkmal **temperature** die empirische Verteilungsfunktion.
4. Erstellen Sie für das Merkmal **hr** ein Histogramm der relativen Häufigkeiten. Nehmen Sie dazu äquidistante Klassen der Breite 5 von 55 bis 90.
5. Erstellen Sie für das Merkmal **temperature** ein Dichtehistogramm. Nehmen Sie dazu äquidistante Klassen der Breite 0.5 von 96 bis 99.5 bzw. 101.

Hinweis: Ein Dichtehistogramm (auch „flächentreues“ Histogramm genannt) ist ein Histogramm, dessen Fläche gleich 1 ist.

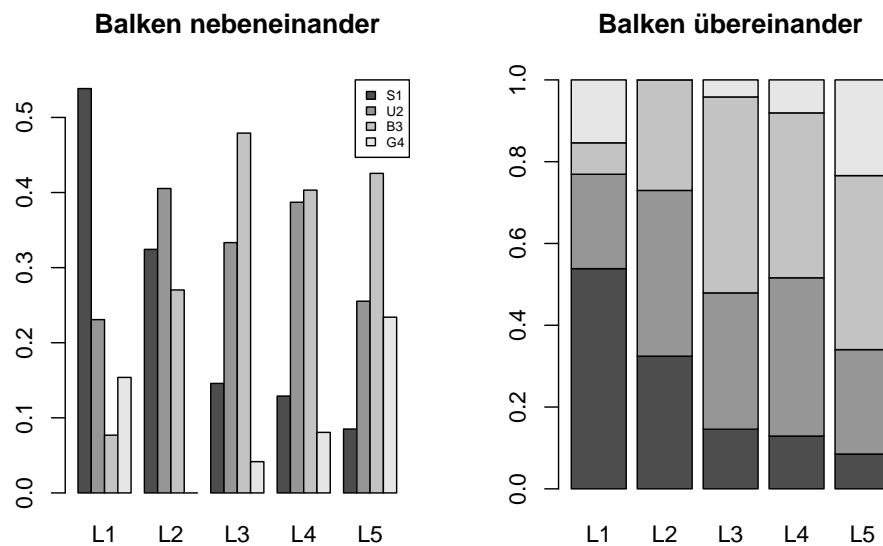
6. Erstellen Sie auf Basis der Klasseneinteilung von Bsp-4 bzw. Bsp-5 das Summenpolygon für **hr** bzw. **temperature**.

Lösungen zum 2. Blatt

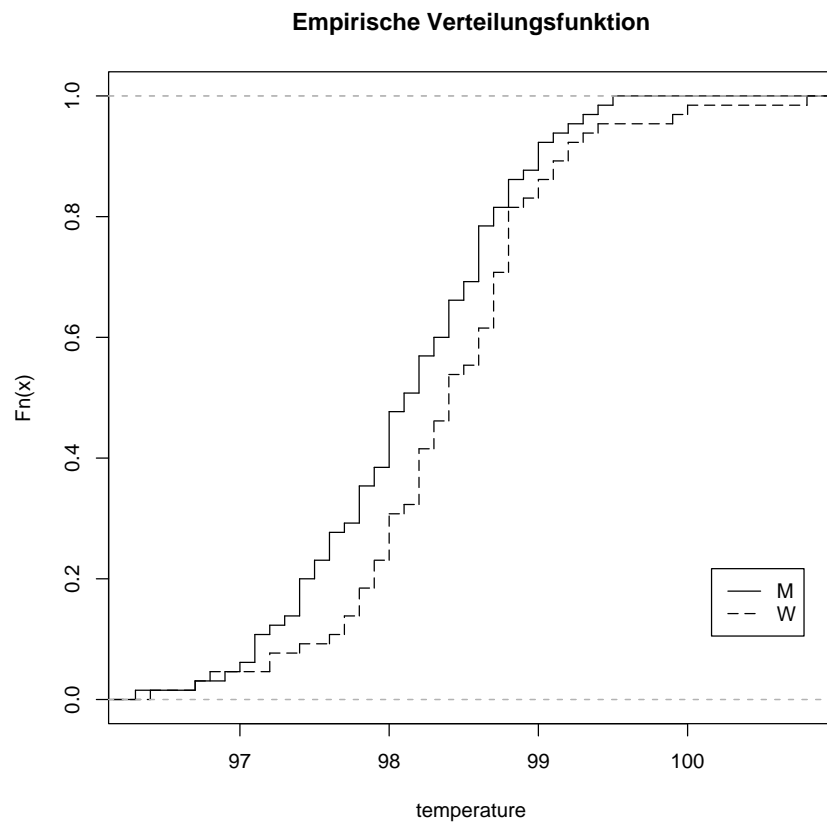
1. UE-Noten: qualitativ, diskret, ordinal
 UE-Leiter: nominal
 (Absolute) Häufigkeiten: Zählvariable, quantitativ, diskret, verhältnisskaliert
temperature: quantitativ, stetig, intervallskaliert
hr: Zählvariable, quantitativ, diskret, verhältnisskaliert
gender: nominal
2. (a) Gesamtbenotung (relative Häufigkeiten):



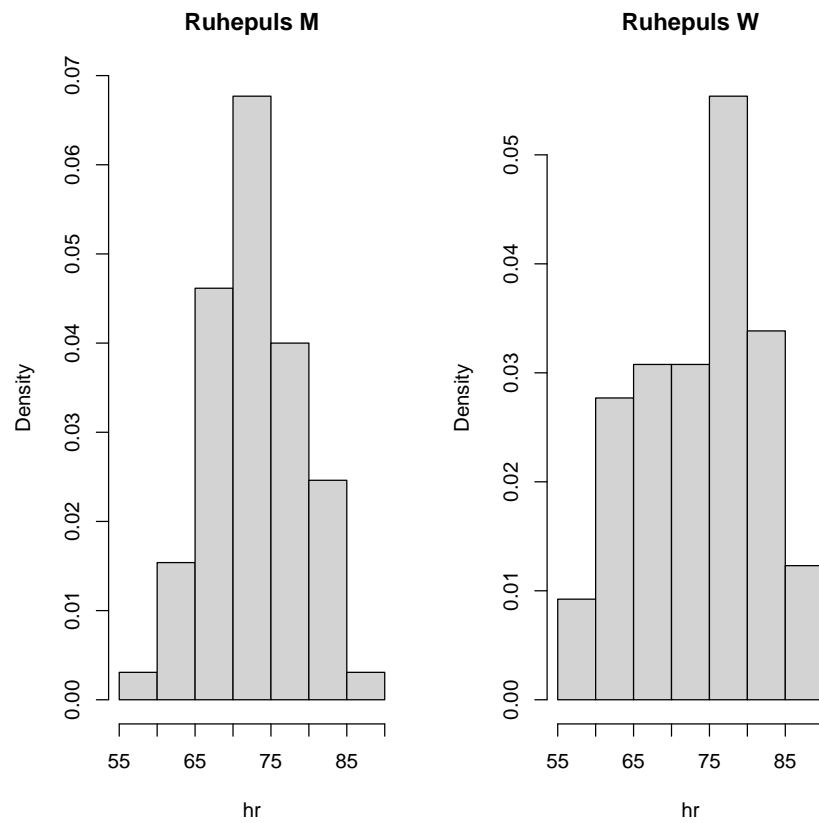
- (b) Nach den UE-Leitern (relative Häufigkeiten):



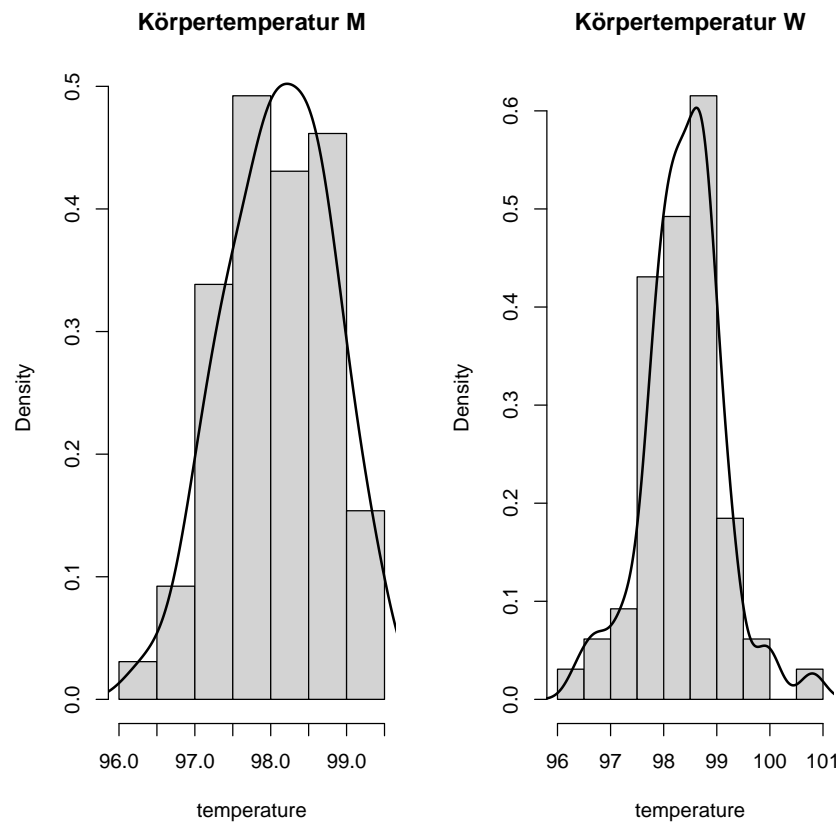
3. Empirische Verteilungsfunktion der Körpertemperatur:



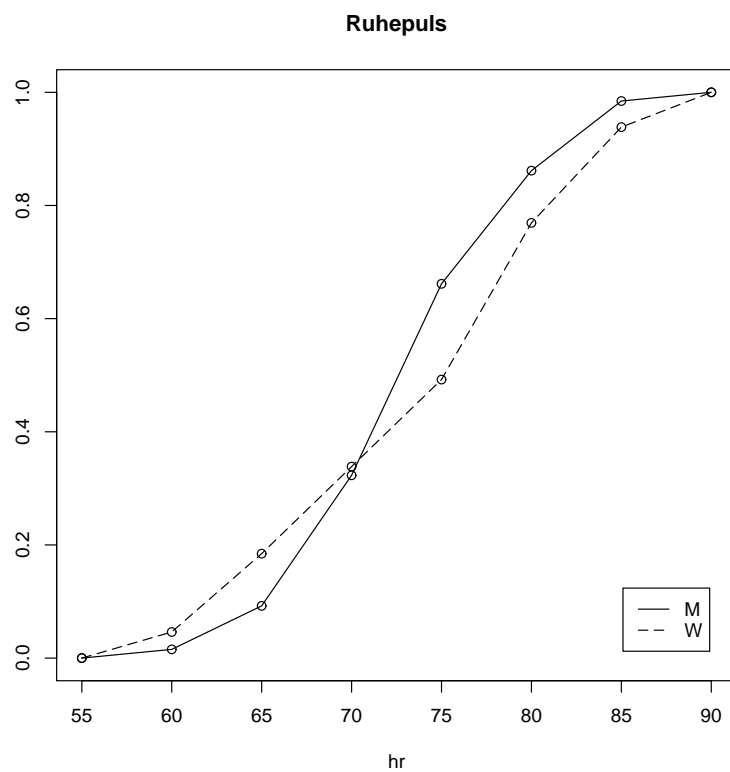
4. (Dichte-) Histogramm des Ruhepulses:



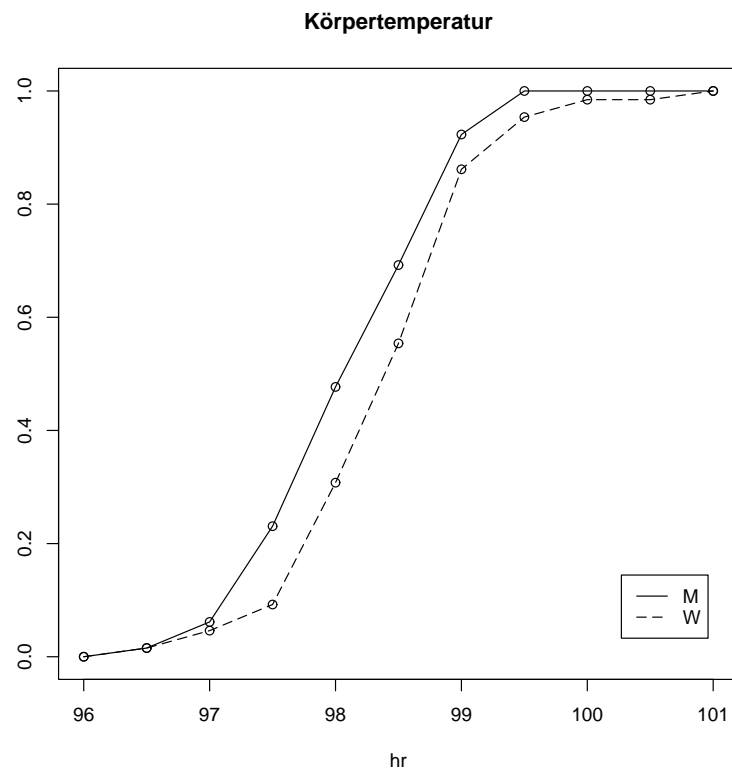
5. (Dichte-) Histogramm der Körpertemperatur (überlagert mit einer „glatten“ Dichteschätzung):



6. Summenpolygon für den Ruhepuls:



Summenpolygon für die Körpertemperatur:



Bem.: Es gibt noch zahlreiche andere Möglichkeiten zur graphischen Aufbereitung von Datensätzen, beispielsweise die Stamm-und-Blatt-Darstellung (eine Art Histogramm mit von den Daten selbst bestimmter Klasseneinteilung). Für die Körpertemperatur ist diese Darstellung gegeben wie folgt (*dos à dos* von M und W):

Körpertemperatur

M	W
3 96	4
97 96	78
4444321110 97	224
998888766655 97	677888999
444433222211000000 98	0000012222233344444
98887766666655 98	5666677777788888889
4321000 99	00112234
5 99	9
100	0
100	8