

Datorövning 1



uwe.menzel@math.uu.se

1. Medelvärde, Median, Standardavvikelse osv.



- Data: <http://www.math.uu.se/~uwe/>
 - grupp1.txt grupp2.txt (tips: "rename worksheet")
- Räkna ut för en grupp: medelvärde, median, kvartiler Q1 och Q3, variationsbredd, mellankvartilvariation, stickprovsvarians och –standardavvikelse
- Ange procentsatsen av data i intervallet $\bar{x} \pm 2s$
- S_{xx} är en storhet som behövs i många sammanhang. Beräkna S_{xx} . Använd nedanstående formel!

$$s^2 = \frac{1}{n-1} S_{xx}$$

2. Korrelation



- Sjukdomar hos mjölkkor som minskar i frekvens med besättningens storlek 2007/08
 - SCB: <http://www.pubkat.scb.se/ViewInfo.aspx?publobjid=8568&lang=SV>
 - Kalvningsförklarning
 - Ladda ner filen "kor.txt"
- Beräkna kovarians och korrelation
- Jämför kovariansmatrisen med varianserna för båda kolumner
- Rita ett spridningsdiagram (scatterplot)
- Finns en korrelation?
- Kan det finnas ett orsakssammanhang också?

3. Rank-korrelation



- File "meander.txt"
- Ranka dessa värden! (tips: Data/Rank)
- Beräkna rank-korrelationskoefficienten! (*Spearman's rank correlation coefficient*)
- Finns en korrelation?
- Beräkna korrelationen för de ursprungliga data (*Pearsons correlation coefficient*) och jämför!

4. Grafik



- Fil "kroppstemp.txt"
- Rita ett histogram över "Age" och över "Temp" (snygga gärna upp dem: färg, text osv.)
- Rita ett histogram med arean normerad till 1
- Rita ett spridningsdiagram "Temp" vs "Age"
 - tror du att de är korrelerade ?
- Rita samma diagram, men gruppera efter "Gender"
- Rita samma diagram, men gruppera efter "Smoke"
- Rita ett lådagram över "Temp"
 - Är fördelningen skev? Jämför med histogrammet!
 - Vilket numeriskt värde har outliern ? Hur gammal är personen som är outlier?

5a. Binomialfördelningen



- En fröpåse innehåller 25 frön. Grobarheten är bara 35%. (Grobarheten är slh. att ett frö utvecklas.)
- Hur stor är sannolikheten att exakt 15 frön groor ? (Calc / Probability distribution ...)
- Vilket antal utvecklade frön skulle du ungefär förvänta dig? Hur stor är sannolikheten för detta antal?
- Låt X vara antalet frön ur denna påse som groor. Vilken fördelning har X ? $X \sim ??$
- Rita sannolikhetsfunktionen för X . (Graph / Probability Distrib. Plot ...)
- Vad är sannolikheten att mindre än 7 frön groor? (Calc / Prob. distrib. ...)
- Vad är sannolikheten att minst 7 frön groor?

5b. Binomialfördelningen



- En säljare säger att grobarheten av en viss sort frö är genomsnittligt 80%. Du köper 20 stycken och 12 av dem groor. Du reklamerar eftersom värdet 80% kan enligt din uppfattning inte stämma. Han menar dock att det alltid finns en viss spridning och att du fick en dålig påse pga. slump.
- Hur stor är sannolikheten att 12 eller mindre än 12 frön groor om grobarheten verkligen är 80%?

Resultat



Resultat 1



	grupp1.txt	grupp2.txt
aritmet. medelvärde	98.04	103.6
median	89	95
kvartil Q_1	75	81
kvartil Q_3	110	120
variationsbredd	120	150
mellankvartilvariation	35	39
varians s^2	836.37	838.68
standardavvikelse s	28.92	28.96

$\bar{x} \pm 2s = (40.2; 155.88)$ $\bar{x} \pm 2s = (45.68; 161.52)$
 $79/83 = 95\%$ $93/97 = 96\%$

Computer_1.MPJ

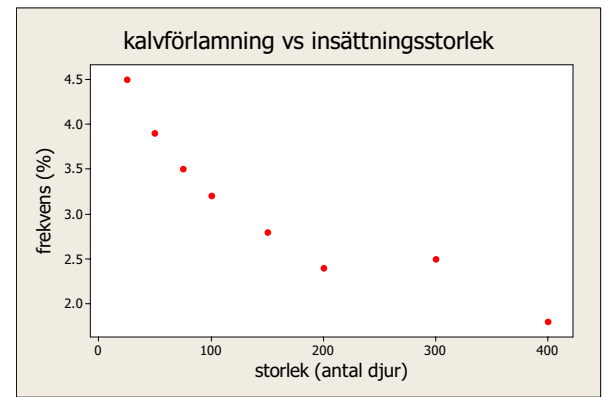
Resultat 2



Pearson correlation: **$r = -0.917$**
P-Value = 0.001

Kovariansmatris:

	s	f
s	17142	-105
f	-105	0.771



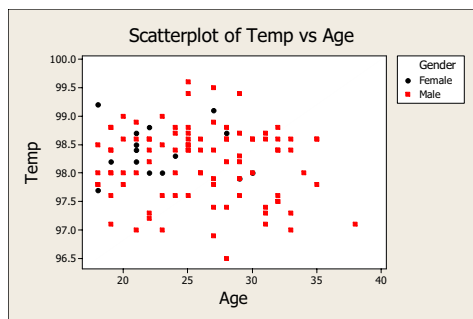
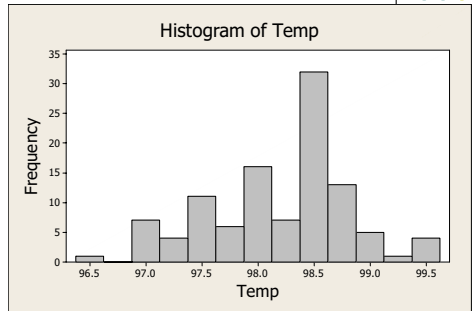
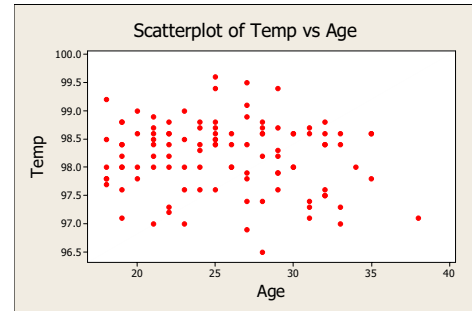
Resultat 3



Spearman's rank-correlation: **$r = 0.967$**
Pearson correlation coefficient: **$r = 0.977$**

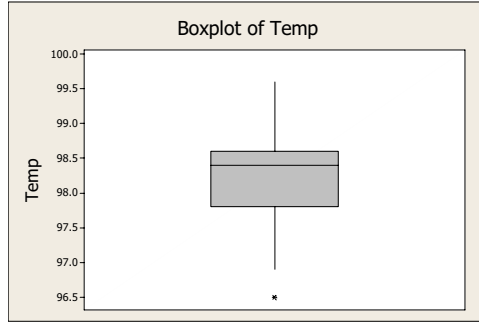
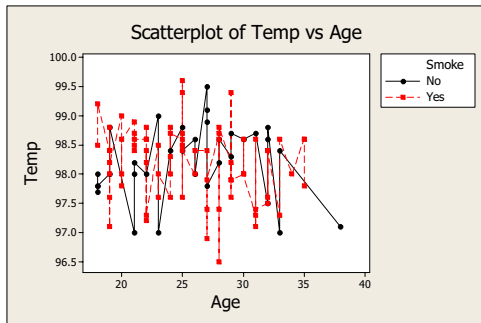
För data som är definerade på en ordinalskala (differenser saknar betydelse, se exemplet "T-tröjor") används rank-korrelationen.

Resultat 4





Resultat 4



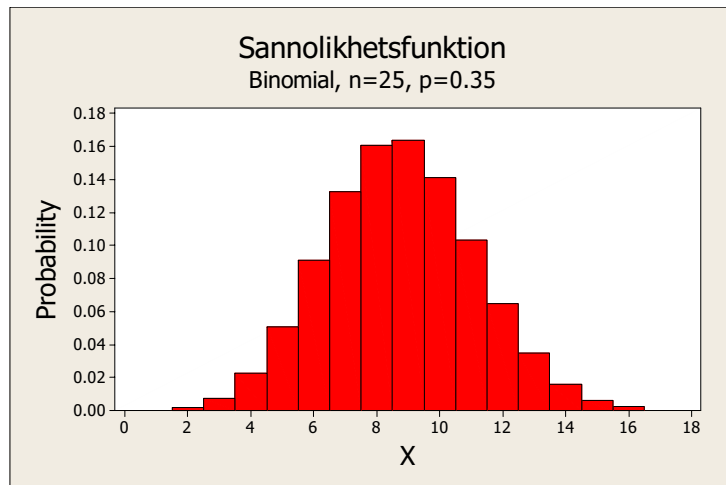
Resultat 5a

- Slh. att exakt 15 frön gror: $p_x(15) = 0.00637$
- Vilket antal utvecklade frön skulle du ungefär förvänta dig? $25 \cdot 0.35 \sim 9$ $p_x(9) = 16.3\%$
- Vilken fördelning har X? $X \sim \text{Bin}(25, 0.35)$
- Sannolikhetsfunktionen för X: nästa sida
- Slh. att mindre än 7 frön gror? $P(X \leq 6) = 0.1734$
- Vad är sannolikheten att minst 7 frön gror?

$$P(X \geq 7) = 1 - P(X \leq 6) = 1 - 0.1734 = 0.8266$$



Resultat 5a



Resultat 5b

- $X \sim \text{Bin}(20, 0.8)$
- $P(X \leq 12) = 0.032$ (Calc / Probability distributions / Binomial / Cumulative Probability ...)
- Slh. är bara 3% att bara 12 (eller mindre) utav 20 frön gror om grobarheten är verkligen 80%. Köpa någon annanstans nästa gång?