

## Ma2c - Prövning nr. 9 (av 9) för betyget E - Sannolikhet och Statistik

*Hjälpmedel : Papper, penna och sudd*

*Obs! Minsta slarvfel kan ge underkänt. Nytt försök tidigast om en vecka.*

Frekvens för en observation är det antal gånger observationen (till exempel ett värde) förekommer i en datamängd.

Relativ frekvens är frekvens för en viss observation (ett visst värde) dividerat med totalt antal observationer (värden).

Man delar ofta in observationerna i klasser och redovisar i histogram. När man samlar in data kallar man hela gruppen man undersöker population och gör antingen en totalundersökning av alla eller en stickprovsundersökning, dvs undersöker en del av populationen.

Medelvärde

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \cdots + x_n}{n}$$

Median är det mittersta värdet då observationerna är sorterade i storleksordning. Om antalet observationer är jämnt är medianen medelvärdet av de två mittersta.

Typvärde är det värde som förekommer flest antal gånger, dvs har högst frekvens.

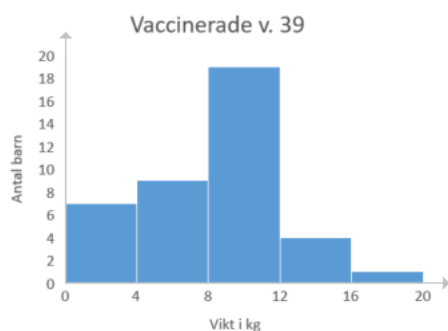
Variationsbredd är skillnaden mellan det största och det minsta värdet i en given datamängd.

Standardavvikelse är ett mått på spridningen av värdena i en datamängd. Den beräknas enligt en formel (som ligger över E-nivå) men erhålls vanligen från digitala hjälpmedel.

Normalfördelat är ett statistiskt material om det fördelar sig enligt en normalfördelningskurva. En sådan har observationsvärden längs x-axeln och observationernas (värdenas) frekvens längs y-axeln. Se Ex. 2 för beskrivning och användning av en normalfördelning.

Regressionsanalys innebär att man anpassar funktioner till mätvärden, till exempel erhållna från en statistisk undersökning. Detta illustreras i Ex. 3.

## Histogram



Ett histogram visar frekvensen av mätvärden i intervall. Exemplet ovan visar hur många barn som vaccinerats vid en klinik under vecka 39 (1989) uppdelade i vikt-klasser 0 – 4 kg, 4 – 8 kg, osv.

## Låddiagram (lådagram)



Ett låddiagram visar med hjälp av medianen samt de så kallade nedre och övre kvartilerna hur spridda värdena i en datamängd är. 25 % av värdena ligger mellan min och  $q_1$  (nedre kvartil), 25 % mellan  $q_1$  och medianen, 25 % mellan medianen och  $q_2$  (övre kvartil) och 25 % mellan  $q_2$  och max.

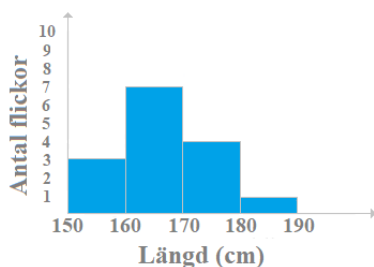
Skriv av följande exempel:

Ex.1 Man mätte längden på flickorna i en klass och fick följande värden (cm):  
173 159 182 178 156 166 166 177 169 162 169 158 171 166 163

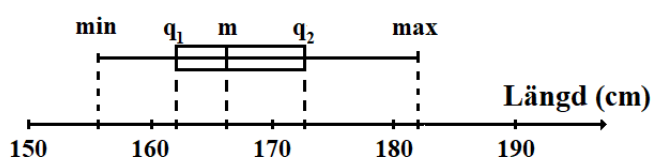
- Ange frekvensen för värdet 169
- Ange relativa frekvensen för värdet 169
- Bestäm variationsbredd
- Bestäm typvärde
- Bestäm median
- Bestäm medelvärde och gör en rimlighetsuppskattning
- Bestäm standardavvikelsen (frivillig uppgift)
- Dela in värdena i fyra lämpliga grupper och rita ett histogram över dem
- Bestäm  $q_1$  och  $q_2$  och rita ett låddiagram som visar värdenas spridning

Lösning:

- a) Värdet 169 förekommer två gånger  $\Rightarrow$  frekvensen är 2
- b) Det är 15 värden  $\Rightarrow$  den relativa frekvensen för 169 är  $2/15 \approx 0.13$
- c) Variationsbredden blir  $(182 - 156) \text{ cm} = 26 \text{ cm}$   
För att fortsätta skriver man värdena i storleksordning (cm):  
156 158 159 162 163 166 166 166 169 169 171 173 177 178 182
- d) Typvärdet är 166 cm
- e) Medianen är 166 cm
- f) Medelvärdet beräknas som summan (2515) av värdena ovan dividerat med antalet (15) till 168 cm (167.67 cm) vilket anses rimligt
- g) Standardavvikelsen beräknas med digitalt hjälpmedel till 7.6 cm (7.603 cm)
- h) Dela in värdena i grupperna I(151–160) cm, II(161–170) cm, III(171–180) cm och IV(181–190) cm. I grupp I finns tre värden, i grupp II sju, i grupp III fyra och i grupp IV ett. Detta resulterar i följande histogram:



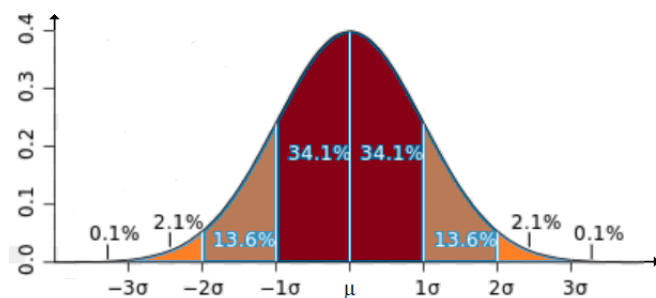
- i) Med  $q_1 = 162 \text{ cm}$  och  $q_2 = 173 \text{ cm}$  ligger 25 % av värdena i varje intervall. Då erhålls lådagrammet:



Ex.2 En normalfördelning visar i procent hur stor del av de observerade värdena som ligger i intervallen:

- mindre än medelvärdet  $\mu$  minus tre gånger standardavvikelsen  $\sigma$
- mellan  $\mu$  minus  $3\sigma$  och  $\mu$  minus  $2\sigma$
- mellan  $\mu$  minus  $2\sigma$  och  $\mu$  minus  $\sigma$
- mellan  $\mu$  minus  $\sigma$  och  $\mu$
- mellan  $\mu$  och  $\mu$  plus  $\sigma$
- mellan  $\mu$  plus  $\sigma$  och  $\mu$  plus  $2\sigma$
- mellan  $\mu$  plus  $2\sigma$  och  $\mu$  plus  $3\sigma$
- större än  $\mu$  plus  $3\sigma$

enligt figuren:



För en grupp på 10000 personer gällde att längden var normalfördelad med medelvärdet 180 cm och standardavvikelsen 5 cm. Besvara följande frågor med hjälp av figuren

- Hur många av personerna var mellan 175 cm och 185 cm långa?
- Hur många av personerna var över 190 cm långa?
- Hur många av personerna var mellan 165 cm och 175 cm långa?
- Hur förändras kurvan om medelvärdet hade varit 170 cm?

Lösning:

- Två gånger 34.1 % av 10000 var mellan 175 cm och 185 cm långa. Detta blir:  
 $68.2\% \text{ av } 10000 = 0.682 \cdot 10000 = 6820 \text{ st.}$
- 2.1 % + 0.1 % av 10000 var över 190 cm långa. Detta blir:  
 $2.2\% \text{ av } 10000 = 0.022 \cdot 10000 = 220 \text{ st.}$
- 2.1 % + 13.6 % av 10000 var mellan 165 cm och 175 cm långa. Detta blir:  
 $15.7\% \text{ av } 10000 = 0.157 \cdot 10000 = 1570 \text{ st.}$
- Om medelvärdet hade varit 170 cm skulle kurvan ligga 10 cm längre åt vänster, det vill säga  $\mu$  skulle ligga där det står  $-2\sigma$  i figuren ovan.

Ex.3 Sambandet mellan hastighet och bromssträcka för en Honda Varadero ges av:

30	60	90	120	150	180 km/h
6	14	29	43	61	83 m

- Bestäm  $a$ ,  $b$  och  $c$  med hjälp t.ex. Excel så att  
 $y = ax^2 + bx + c$  blir en anpassning till mätvärdena.
- Beräkna med  $y$  enligt a) den hastighet som ger en bromssträcka på 50 m.
- Beräkna med  $y$  enligt a) den bromssträcka som ges av en hastighet på 110 km/h.

Lösning:

- a) I ett digitalt hjälpmedel, t.ex. Excel, för man in hastigheterna i kolumn A och bromssträckorna i kolumn B, märker dessa kolumner och väljer 'Infoga punktdiagram'. Sedan klickar man på det stora plusstecknet, väljer 'Trendlinje' samt 'Fler alternativ...' vid högerpilen. Till höger klickar man så i 'Polynom' och 'Visa ekvation i diagramet'. Detta ger  $y = 0.0016x^2 + 0.1726x - 1$ .
- b) Sätt  $y = 50$  och lös ekvationen  $50 = 0.0016x^2 + 0.1726x - 1 \Rightarrow 0.0016x^2 + 0.1726x - 51 = 0$   
Rita kurvan i GeoGebra  $y_1 = 0.0016x^2 + 0.1726x - 51$  och läs av var den skär x-axeln  $\Rightarrow x \approx 133 \text{ km/h}$
- c) Sätt  $x = 110 \Rightarrow y = 0.0016 \cdot 110^2 + 0.1726 \cdot 110 - 1 \approx 37 \text{ m}$

Redovisa fullständiga, korrekta lösningar av följande uppgifter för betyget E:

- Man mätte längden på flickorna i ett basketlag och fick följande värden (cm):  
188 174 197 193 171 181 181 192 184 177 184 173 186 181 178
  - Ange frekvensen för värdet 184
  - Ange relativa frekvensen för värdet 184
  - Bestäm variationsbredd
  - Bestäm typvärde
  - Bestäm median
  - Bestäm medelvärde och gör en rimlighetsuppskattning
  - Bestäm standardavvikelsen (frivillig uppgift)
  - Dela in värdena i fyra lämpliga grupper och rita ett histogram över dem
  - Bestäm  $q_1$  och  $q_2$  och rita ett lådagran som visar värdenas spridning
- För en grupp på 1000 personer gällde att vikten var normalfördelad med medelvärdet 100 kg och standardavvikelsen 20 kg. Besvara följande frågor med hjälp av figuren i Ex. 2
  - Hur många av personerna vägde mellan 80 kg och 100 kg?
  - Hur många av personerna vägde över 120 kg?
  - Hur många av personerna vägde mellan 40 kg och 80 kg?
  - Hur förändras kurvan om medelvärdet hade varit 110 kg?

3. Sambandet mellan hastighet och bromssträcka för en viss scouter ges av följande:

25	50	75	100	125	150 km/h
5	12	29	52	90	120 m

- a) Bestäm  $a, b$  och  $c$  med hjälp t.ex. Excel så att  
 $y = ax^2 + bx + c$  blir en anpassning till mätvärdena.
- b) Beräkna med  $y$  enligt a) den hastighet som ger en bromssträcka på 40 m.
- c) Beräkna med  $y$  enligt a) den bromssträcka som ges av en hastighet på 90 km/h.