

Test for normalfordeling i WordMat

Indhold

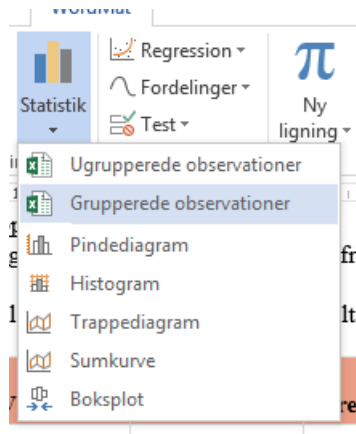
Grupperede observationer	1
Ugrupperede observationer	3
Analyse af normalfordelt observationsæt	4

Grupperede observationer

Vi tager udgangspunkt i et eksempel fra *MAT B2 stx*, s. 31 (p418 i iBogen), hvor der er en tabel over 600 personers vægt, fordelt i intervaller á 5 kg:

Vægt, kg	Antal personer	Frekvens i %	Kumuleret frekvens, %
]40; 45]	6	1,0	1,0
]45; 50]	8	1,3	2,3
]50; 55]	28	4,7	7,0
]55; 60]	54	9,0	16,0
]60; 65]	92	15,3	31,3
]65; 70]	114	19,0	50,3
]70; 75]	122	20,3	70,6
]75; 80]	90	15,0	85,6
]80; 85]	55	9,2	94,8
]85; 90]	22	3,7	98,5
]90; 95]	5	0,8	99,3
]95; 100]	4	0,7	100,0

Åbn WordMats Excel-ark til behandling af grupperede observationer under
 WordMat → Graftegning → Statistik → Grupperede observationer
 (På Mac skal du vælge det eneste Excel-dokument, der er at vælge imellem i den sidste menu.)



Excel-arket popper ikke op. Du skal selv finde det i oversigten over åbne programmer (i Windows som regel i den nederste vandrette bjælke på skærmen).

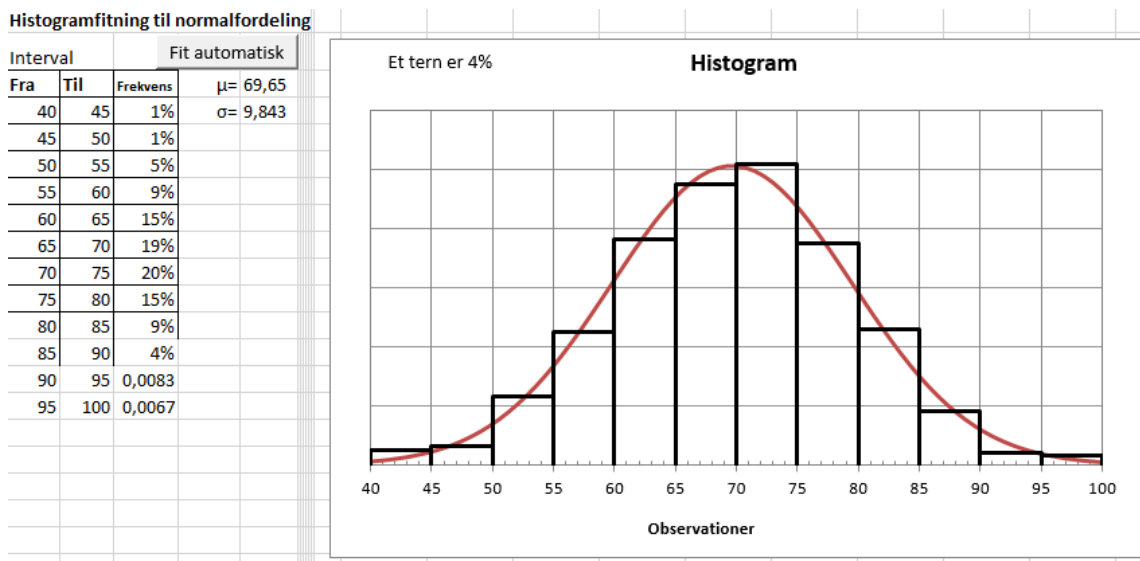
I Excel-arket: Gå ind under fanen "Data". Klargør intervallerne ved at indtaste nederste intervalendepunkt, øverste intervalendepunkt og intervallængden øverst på arket:

	start	slut	Interval længde
	40	100	5

Nu står intervallerne klar nedenunder. Skriv eller kopier hyppighederne (antal personer i dette eksempel) ind og klik på "Kopier til øvrige ark".

Gå ind under fanen "Histogram-Fit". Her ses histogrammet for observationerne.

WordMat har desuden beregnet middelværdien $\mu = 69,65$ og spredningen $\sigma(X) = 9,843$ og ud fra disse to værdier tegnet den klokkeformede frekvensfunktion for en normalfordeling:



Vi kan se, at histogrammet passer nogenlunde til frekvensfunktionen for en normalfordeling. Derfor må der være tale om et tilnærmelsesvis normalfordelt observationsæt.

Bemærk, at WordMats beregning af μ og $\sigma(X)$ afhænger af, hvordan materialet er grupperet. Hvis der er tale om ugrupperede observationer, fås derfor et mere præcist resultat ved at undlade at gruppere observationerne og i stedet følge vejledningen nedenfor.

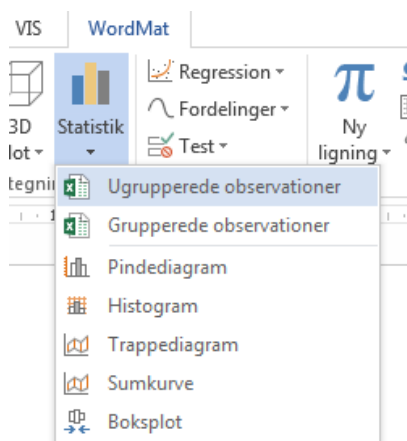
Ugrupperede observationer

30 voksne mænds højde er målt i cm:

189	173	165	169
167	174	173	186
164	178	182	173
160	190	179	188
153	177	166	175
181	157	183	176
176	172	174	
179	180	186	

Vi ønsker at undersøge, om højderne er tilnærmelsesvis normalfordelt og i så fald at bestemme middelværdien μ og spredningen $\sigma(X)$.

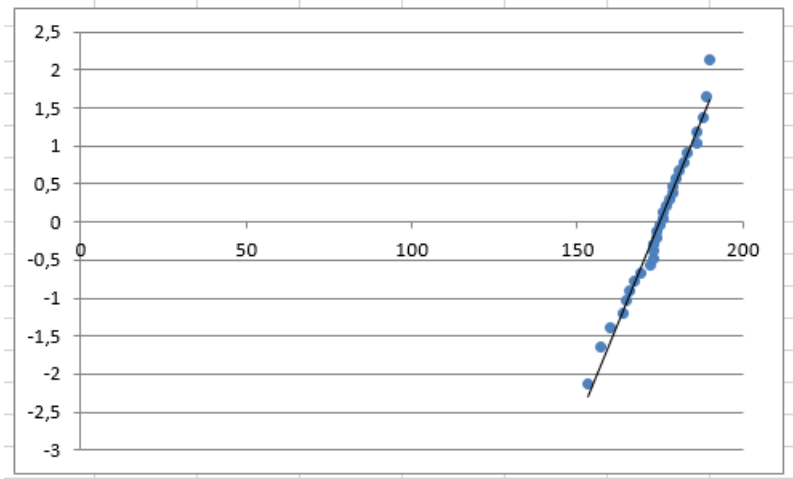
Åbn WordMats Excel-ark til behandling af ugrupperede observationer under
WordMat \rightarrow Graftegning \rightarrow Statistik \rightarrow Ugrupperede observationer
(På Mac skal du vælge det eneste Excel-dokument, der er at vælge imellem i den sidste menu.)



Excel-arket popper ikke op. Du skal selv finde det i oversigten over åbne programmer (i Windows som regel i den nederste vandrette bjælke på skærmen).

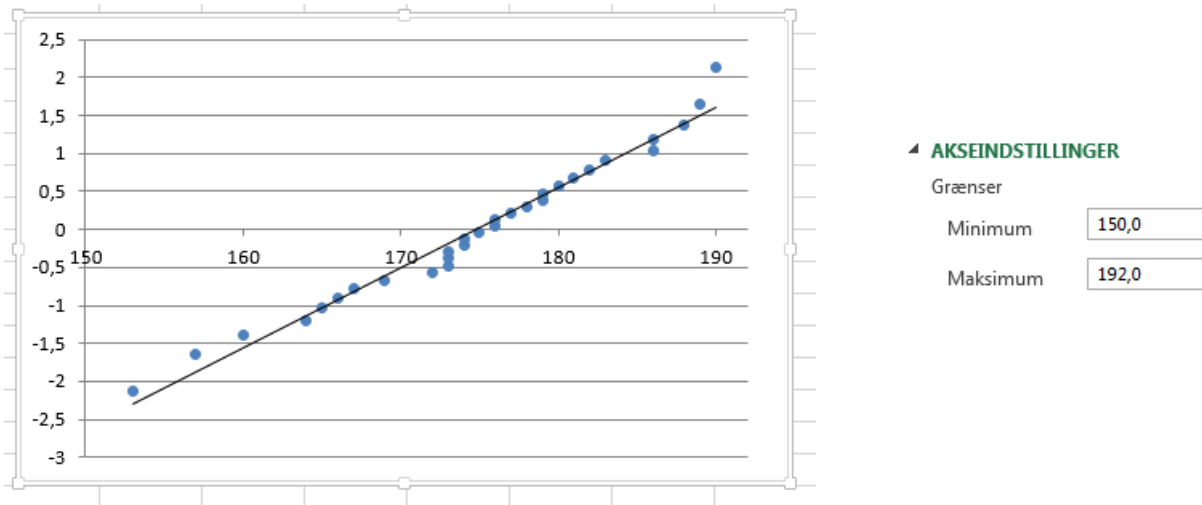
I Excel-arket: Gå ind under fanen "Data". Skriv eller kopier observationerne ind under "Rå data" og klik på "Kopier til øvrige ark".

Gå ind under fanen "Normal-Plot". WordMat har her lavet et normalfordelingsplot af de kumulerede frekvenser og tegnet den bedste rette linje igennem punkterne:



Op ad 2.-aksen er angivet afvigelsen fra middelværdien μ . Fx angiver -1 på 2.-aksen, at her er tale om $\mu - 1 \cdot \sigma(X)$.

Hvis punkterne tilnærmelsesvis ligger på en ret linje, er der tale om et tilnærmelsesvis normalfordelt observationsæt. Det ser ud til at være tilfældet, men for at lave vurderingen ordentligt, skal vi zoome ind på den relevante del af 1.-aksen. Dette gøres ved at højreklikke på 1.-aksen og vælge "Formatér akse". I vinduet, som kommer frem til højre i skærmen, vælges herefter under "Grænser" det relevante interval på 1.-aksen, som får punkterne til at være pænt fordelt i hele udsnittet af koordinatsystemet. I dette tilfælde vil minimum: 150 og maksimum: 192 være passende valg:



Vi kan nu mere sikkert se, at punkterne tilnærmelsesvis ligger på en ret linje, hvormed der er tale om et tilnærmelsesvis normalfordelt observationsæt.

WordMat angiver desuden middelværdi og spredning:

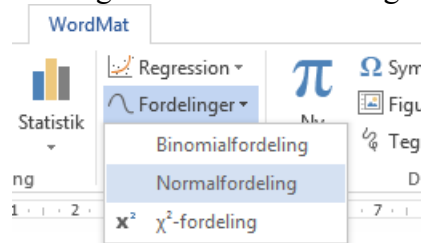
$$\begin{aligned} \mu &= 174,8333 \text{ (Skæring med x-aksen)} \\ \sigma &= 9,456599 \text{ (Invers hældning af linjen)} \end{aligned}$$

Analyse af normalfordelt observationsæt

Når vi har konstateret, at et observationssæt er tilnærmelsesvis normalfordelt, noterer vi middelværdien μ og spredning $\sigma(X)$. Husk så mange decimaler, at du undgår afrundingsfejl.

Analysen laves herefter i WordMats Excel-ark til undersøgelse af normalfordelinger. Dette åbnes under:

WordMat → Statistik/sands. → Fordelinger → Normalfordeling



Excel-arket popper ikke op. Du skal selv finde det i oversigten over åbne programmer (i Windows som regel i den nederste vandrette bjælke på skærmen).

Jeg tager fat på ovenstående eksempel fra afsnittet ”Grupperede observationer”. Her har vi fundet, at

$$\mu = 69,6500000$$

$$\sigma = 9,8426368$$

De to tal indtastes øverst til venstre i Excel-arket:

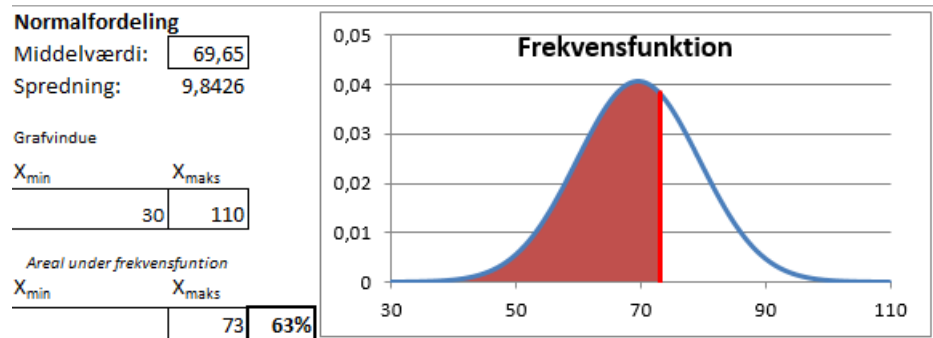
Normalfordeling	
Middelværdi:	69,65
Spredning:	9,8426

Under ”Grafvindue” vælges et passende udsnit af 1.-aksen for de to grafer til højre. Som tommelfingerregel får vi pæne udsnit af kurverne ved at vælge ca. $\mu - 3 \cdot \sigma$ som X_{\min} og ca. $\mu + 3 \cdot \sigma$ som X_{\max} . Det vil i dette eksempel sige:

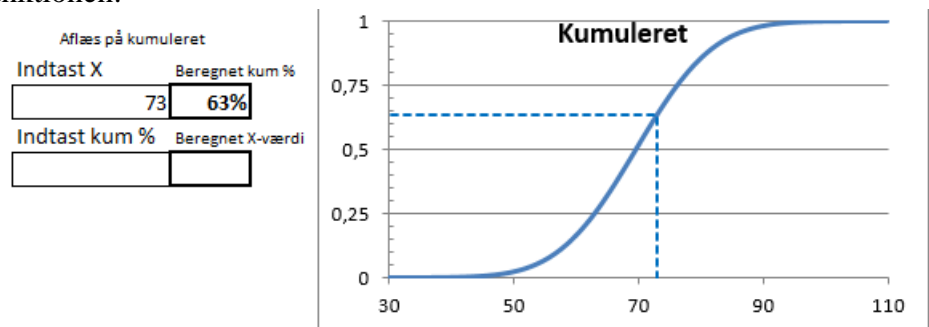
Grafvindue	
X_{\min}	X_{\max}
30	110

Vi er nu klar til at besvare forskellige spørgsmål. Fx:

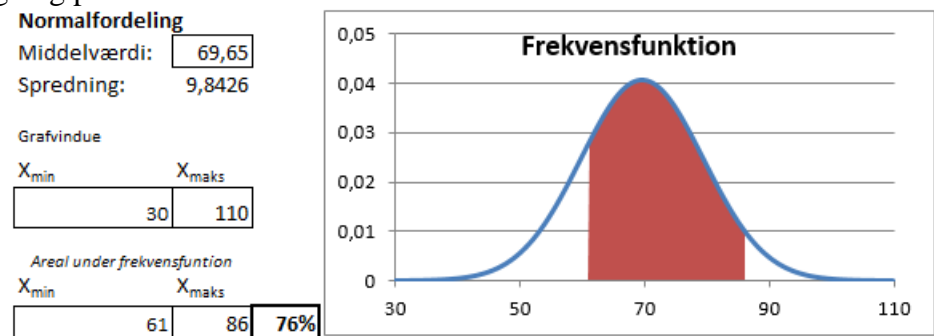
- Hvor mange % vejer 73 kg eller derunder? Kan løses på 2 måder:
 1. Ved at indtaste 73 i feltet X_{\max} under ”Areal under frekvensfunktion”, hvorved WordMat giver løsningen 63% og samtidig laver den relevante indtegning på frekvensfunktionen:



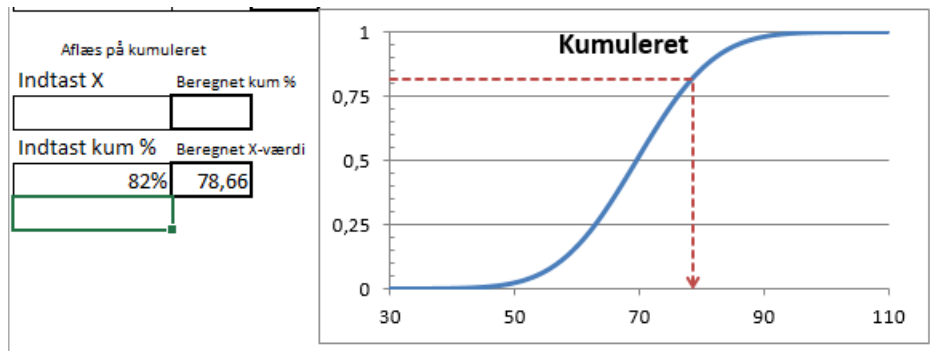
2. Ved at indtaste 73 i feltet "Indtast X" under "Aflæs på kumuleret", hvorved WordMat giver løsningen 63% og samtidig laver den relevante indtegning på fordelingsfunktionen:



- Hvor mange % vejer mellem 61 og 86 kg? Under "Areal under frekvensfunktion": Indtast 61 under X_{\min} og 86 under X_{\max} , hvorved WordMat giver løsningen 76% og samtidig laver den relevante indtegning på frekvensfunktionen:



- Hvad vejer de tungeste 18% af befolkningen? Dette besvares ved at bestemme 82%-fraktilen under "Aflæs på kumuleret". Indtast 82 i feltet "indtast kum %". WordMat giver den beregnede X-værdi 78,66 og laver samtidig den relevante indtegning på fordelingsfunktionen:



Dvs. de tungeste 18% vejer ca. 79 kg eller derover.