**Kvartiler**

Kvartilerna q1, q2 och q3 delar in ett **storleksordnat** datamaterial i fyra lika stora delar, som vardera innehåller 25 % av observationerna.

För ett **ej klassindelat** material är q2 helt enkelt medianen, q1 medianen av undre hälften av observationerna och q3 medianen av övre hälften av observationerna. Om vi har ett udda antal observationer ingår medianen q2 både i undre och övre hälften av observationerna.

För ett **klassindelat** material beräknar man först en kumulativ frekvenstabell. D.v.s. istället för att ange andelen av observationerna i en viss klass så anger man andelen av observationerna i klassen plus alla klasser under.

**Exempel:** Följande är exempel 7 från boken. Data i tabell 2.11 är omräknade till relativa frekvenser.

|  |  |  |
| --- | --- | --- |
| **Ålder** | **Anställda, relativ frekvens, %** | **Kumulerad %** |
| -24 | 3 | 3 |
| 25-29 | 13 | 3+13=16 |
| 30-34 | 29 | 16+29=45 |
| 35-39 | 32 | 77 |
| 40-44 | 16 | 93 |
| 45-49 | 7 | 100 |
| **Totalt** | 100 |  |

För att beräkna en kvartil i det klassindelade materialet tar man först reda på i vilken klass kvartilen ligger och använder följande formel:

**Exempel:** Vi söker q3 i exemplet ovan. Vi ser om vi tittar på kumulerad % i tabellen att q3 måste ligga i klassen 34-39 år eftersom 75 % av data ska vara under q3. Undre klassgräns är 34,5. Klassbredden är 39,5-34,5=5. Kumulerad % i klassen under är kumulerad % i klassen 30-34 år, alltså 45 %. I klassen där q3 ingår finns 32 % av observationerna. Formeln ger då:

Vad gör formeln egentligen? Först har vi ju tagit reda på att q3 är i klassen 34-39 år d.v.s. ligger någonstans mellan 34,5 och 39,5. Vi vet inte var i intervallet. Man kan dock gissa att q3 borde ligga ganska nära 39,5 (varför?[[1]](#footnote-2)). Formeln gör en interpolation. Hur kan man se i följande figur:

I grafen är kumulerad % plottad mot övre klassgräns. Räta linjer är dragna mellan punkterna vilket svarar mot att anta att data är jämt fördelade i varje klass. Vi ska alltså hitta vilken ålder som motsvarar 75 % på den räta linjen.

q3

75 %

Den som är intresserad av ett sätt som formeln kan härledas på kan läsa resten, annars går det utmärkt att hoppa över! Nedan är en uppförstoring av en del av grafen:

I figuren finns två likformiga trianglar. Eftersom de är likformiga förhåller sig deras katetrar till varandra på samma sätt. Alltså är h2/h1=b2/b1. Om vi stoppar in värden på katetrarna får vi:

h1

b1

b2

h2

(34,5;45 %)

(39,5;77 %)

(q3;75 %)

1. Vi vet att 77 % av observationerna är lägre än 39,5. Samtidigt är 45 % av observationerna lägre än 34,5. Alltså borde q3 ligga ganska nära klassens övre gräns. [↑](#footnote-ref-2)