LKT325/LMA521: Faktorförsök

Föreläsning 2

Anders Hildeman

Innehåll

- Referensfördelning
- Referensintervall
- Skatta variansen
  1. Flera mätningar i varje grupp.
  2. Antag att vissa effekter inte existerar
  3. Normalfördelningspapper

Referensfördelning

- Hittills har vi bara brytt oss om vår skattning av väntevärdena. Vi vet sedan tidigare i kursen att skattningen av väntevärdet i alla praktiska tillämpningar kommer skilja sig från det sanna väntevärdet pga slumpen.
- Detta innebär bl.a. att faktorer som egentligen inte har någon påverkan kommer ha en skattad effekt som är skild från noll.
- Hur skall vi veta om vår skattade effekt är en verklig effekt eller bara mätbrus?

- Tanken är att vi antar att vi känner till vilken typ av sannolikhetsfördelning som mätbruset följer.
- I den här kursen kommer vi bara anta att mätbruset är normalfördelat. Detta är ofta ett rimligt antagande (men inte alltid).

Referensfördelning

Sannolikhetsfördelningen för en skattad effekt (t.ex. l_a), om faktorn egentligen inte har någon effekt.
Denna sannolikhetsfördelning kan uppkomma på grund av mätbrus i utrustningen eller på grund av störande faktorer som vi inte har någon kontroll över.

- Från början brukar man varken veta vilka faktorer som är signifikanta eller referensfördelningen.
- Om vi antar normalfördelning (och samma varians, σ^2, för alla mätgrupper) så behöver vi bara känna till, σ^2, för att kunna skapa ett konfidensintervall.

- Från början brukar man varken veta vilka faktorer som är signifikanta eller referensfördelningen.
- Om vi antar normalfördelning (och samma varians, σ^2, för alla mätgrupper) så behöver vi bara känna till, σ^2, för att kunna skapa ett konfidensintervall.

$$
\begin{aligned}
Var[l_a] = & Var\left[2\frac{\hat{\mu}_{(+1,-1,\ldots)} + \hat{\mu}_{(+1,+1,\ldots)} + \cdots}{N}\right] \\
& + Var\left[2\frac{\hat{\mu}_{(-1,-1,\ldots)} + \hat{\mu}_{(-1,+1,\ldots)} + \cdots}{N}\right] \\
& = \frac{4}{N^2}\left(\frac{\sigma^2}{n} + \frac{\sigma^2}{n} + \ldots\right) = \frac{4}{N^2}\frac{N\sigma^2}{n} = \frac{4\sigma^2}{Nn},
\end{aligned}
$$

där N är antal olika grupper och n är antal mätningar inom varje grupp.

- Om referensfördelningen är normalfördelad så kommer de skattade effekterna, i de fall då inga riktiga effekter existerar, vara normalfördelade med $N(0, \frac{4\sigma^2}{Nn})$.
- Konfidensintervall kan beräknas såsom vi är vana vid:

$$I_a \pm Z_{\alpha/2}\frac{2\sigma}{\sqrt{Nn}}$$

- Tricket till att avgöra om den skattade effekten skall anses signifikant eller inte är alltså huruvida konfidensintervallet inkluderar 0 eller inte.
  Är 0 inkluderat kan vi inte utesluta att effekten egentligen inte existerar och det bara var "mätbrus" som gav oss ett värde skilt från 0.
- Vi får exakt samma slutsats om vi istället centrerar intervallet i 0 och tittar på vilka skattade effekter som är inkluderat i intervallet eller inte. Vi kallar det då istället för "**referensintervall**".

- Om vi nu inte känner till σ^2, vilket man sällan gör. Hur skall vi då göra för att avgöra vilka effekter som är signifikanta eller inte?
- Låt oss titta på tre olika metoder. Var och en har sitt eget användningsområde.

Metod I: Flera mätningar i varje grupp

- Även om vi inte känner till σ^2 så kan vi skatta den med s^2 (det här känner vi igen från kapitlet om konfidensintervaller).
- Om vi har flera mätningar i varje grupp av faktornivåer: Räkna ut s^2 för varje grupp (låt s_i^2 vara den skattade variansen från mätningarna i grupp i). Det går då att slå ihop skattningarna med formeln:

$$s^2 = \frac{\sum_i (n_i - 1)s_i^2}{\sum_i (n_i - 1)}.$$

- Referensintervallet kan sedan beräknas med hjälp av t-fördelningen såsom vi lärt oss,

$$0 \pm t_{\alpha/2}\left(\nu = \sum_i (n_i - 1)\right) \frac{2s}{\sqrt{Nn}}.$$

Exempel

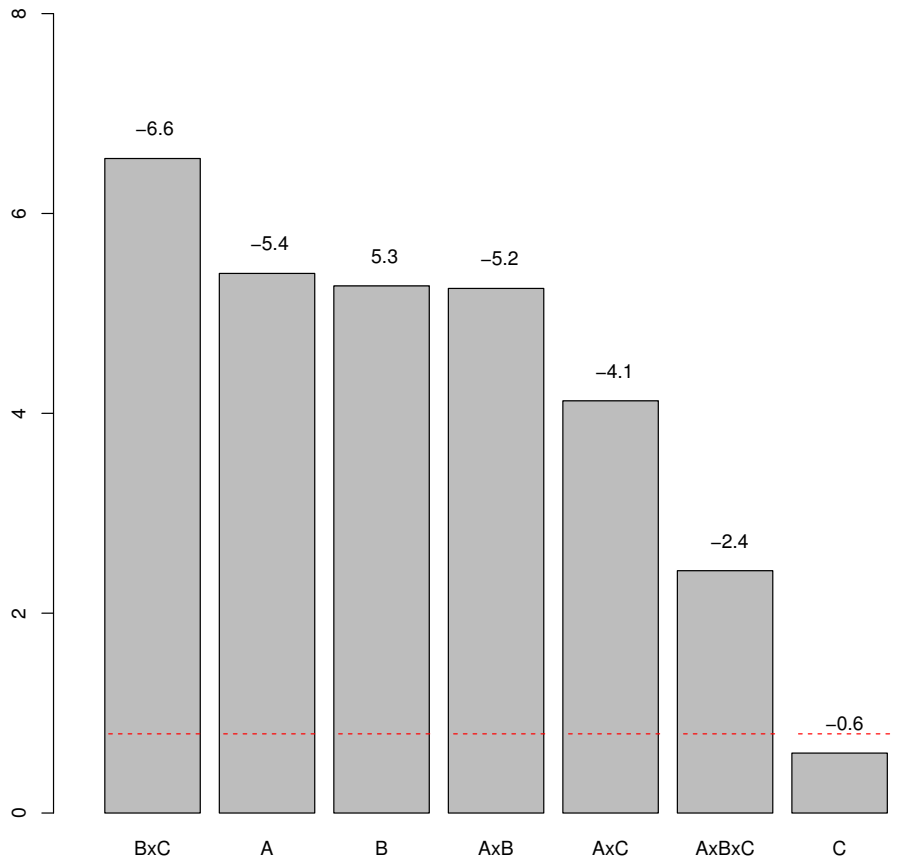
Nr	A	B	C	AB	BC	AC	ABC	Resultat	s^2
1	-	-	-	+	+	+	-	$y_1 = [3.7, 2.8]$	$s_1^2 = 0.405$
2	+	-	-	-	+	-	+	$y_2 = [4.8, 4.8]$	$s_2^2 = 0.0$
3	-	+	-	-	-	+	+	$y_3 = [18.7, 17.1]$	$s_3^2 = 1.28$
4	+	+	-	+	-	-	-	$y_4 = [13.5, 14.1]$	$s_4^2 = 0.18$
5	-	-	+	+	-	-	+	$y_5 = [10.1, 11.7]$	$s_5^2 = 1.28$
6	+	-	+	-	-	+	-	$y_6 = [8.8, 9.3]$	$s_6^2 = 0.125$
7	-	+	+	-	+	-	-	$y_7 = [17.7, 16.9]$	$s_7^2 = 0.320$
8	+	+	+	+	+	+	+	$y_8 = [0.4, -0.2]$	$s_8^2 = 0.18$

Exempel

Nr	A	B	C	AB	BC	AC	ABC	Resultat	s^2
1	-	-	-	+	+	+	-	$y_1 = [3.7, 2.8]$	$s_1^2 = 0.405$
2	+	-	-	-	+	-	+	$y_2 = [4.8, 4.8]$	$s_2^2 = 0.0$
3	-	+	-	-	-	+	+	$y_3 = [18.7, 17.1]$	$s_3^2 = 1.28$
4	+	+	-	+	-	-	-	$y_4 = [13.5, 14.1]$	$s_4^2 = 0.18$
5	-	-	+	+	-	-	+	$y_5 = [10.1, 11.7]$	$s_5^2 = 1.28$
6	+	-	+	-	-	+	-	$y_6 = [8.8, 9.3]$	$s_6^2 = 0.125$
7	-	+	+	-	+	-	-	$y_7 = [17.7, 16.9]$	$s_7^2 = 0.320$
8	+	+	+	+	+	+	+	$y_8 = [0.4, -0.2]$	$s_8^2 = 0.18$

$$s^2 = \frac{1 \cdot s_1^2 + 1 \cdot s_2^2 + \ldots}{8} = \frac{0.405 + 0 + 1.28 + \ldots}{8} = 0.47125$$

$$t_{0.05/2}\,(\nu = 8)\,\frac{2s}{\sqrt{8 \cdot 2}} = 2.306\frac{2 \cdot \sqrt{0.47125}}{\sqrt{16}} = 2.306 \cdot 0.343 = 0.79233$$



Figur: Paretodiagram med referensintervall med skattad s^2.

Metod II: Antag icke-existerande effekter

- Metod I är bäst att använda då vi faktiskt har flera mätningar i varje grupp. Vanligtvis kostar dock mätningar tid och pengar så vad skall man göra då man bara har råd med en mätning per grupp?

Metod II: Antag icke-existerande effekter

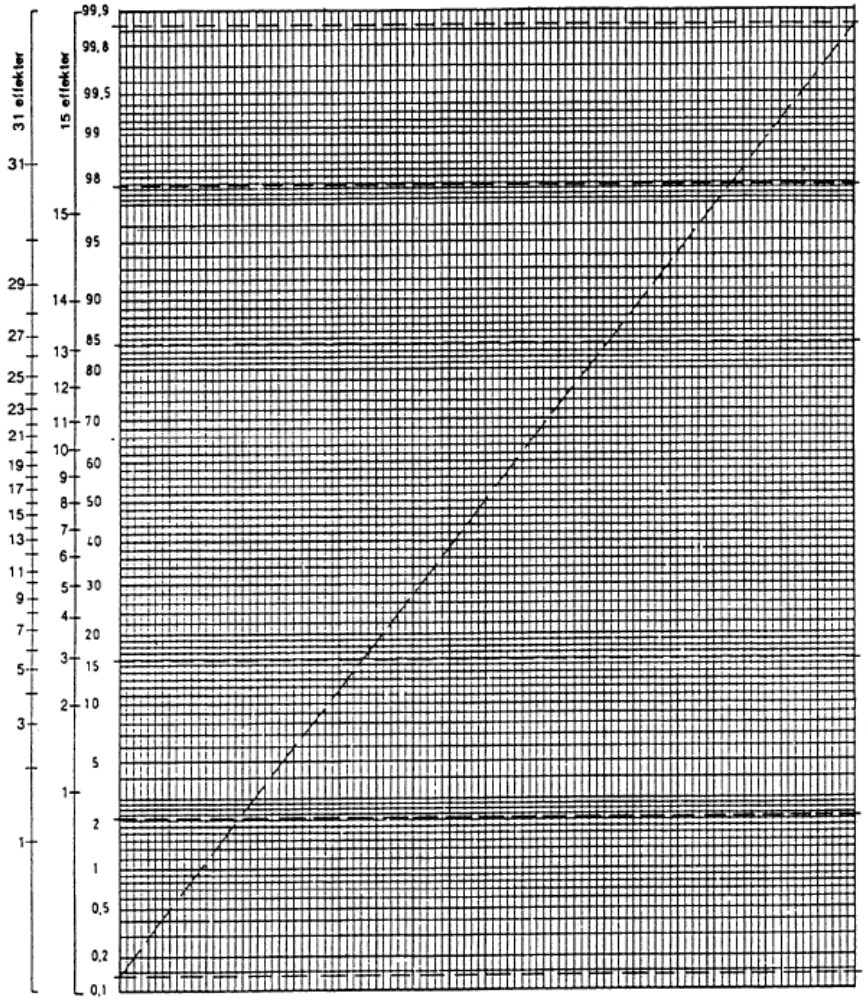
- Metod I är bäst att använda då vi faktiskt har flera mätningar i varje grupp. Vanligtvis kostar dock mätningar tid och pengar så vad skall man göra då man bara har råd med en mätning per grupp?
- Antag att du känner till att några effekter egentligen inte existerar.
  Genom att behandla skattningarna av dessa effekter som olika utfall från referensfördelningen, $N(0, \frac{4\sigma^2}{Nn})$, så kan man sedan skatta σ^2.

$$s^2 = \frac{Nn}{4} \sum_{i \in J} \frac{l_i^2}{N_J},$$

  här är J mängden av grupper som vi inte tror har någon effekt och N_J är antal grupper i J.
- Märk väl att vi inte delar på $N_J - 1$ utan bara på N_J.

Metod III: Normalfördelningspapper

- Den sista varianten är en grafisk metod där vi använder oss av ett s.k. normalfördelningspapper.
- Med hjälp av ett sådant diagram så kan man identifiera ifall mycket av datan ser ut att följa samma normalfördelning. Det går även att se vilka datapunker som inte verkar följa samma fördelning.
- Det här låter ju bra eftersom endast effekter skattade från referensfördelningen borde följa samma normalfördelning. Alla andra effekter borde inte se ut att komma från en gemensam fördelning



(a) Normalfördelningspapper

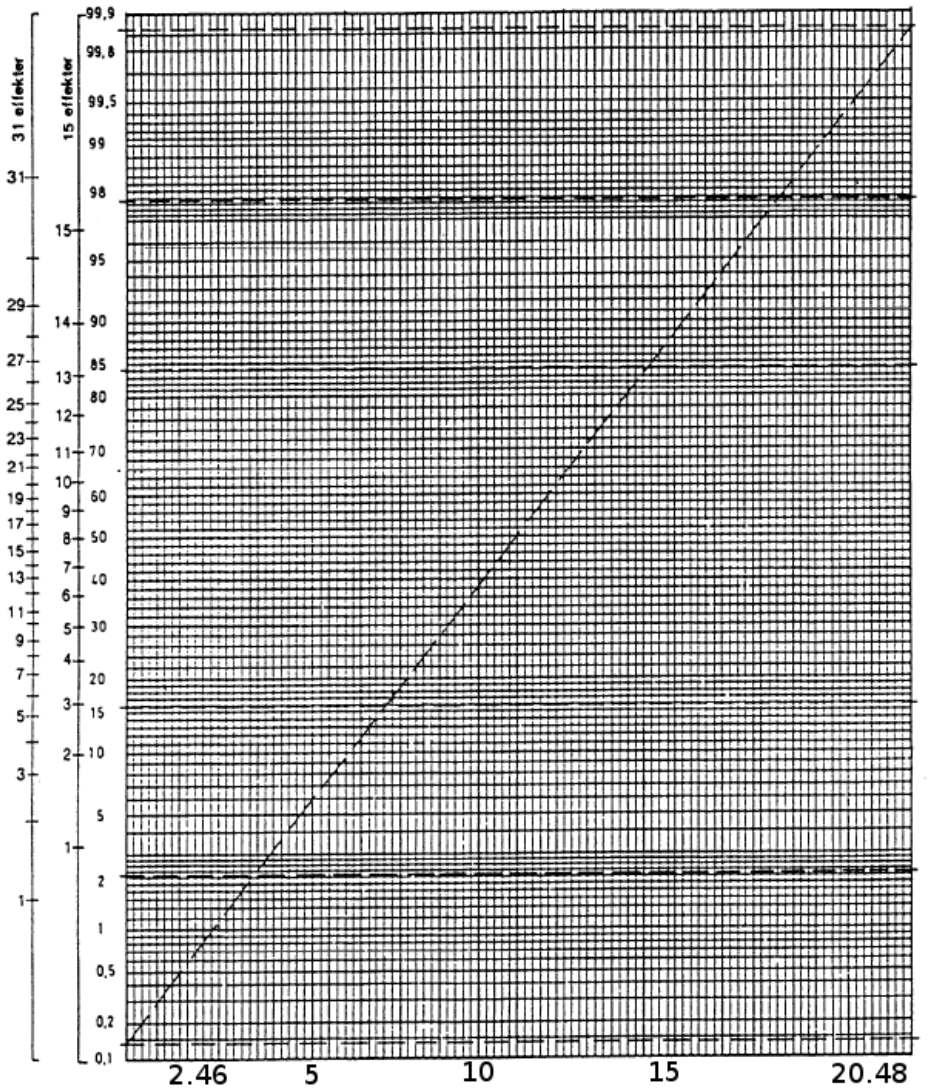
Diagram där man på y-axeln har skrivit ut sannolikheter med ett avstånd emellan som kommer ge en rak linje för en normalfördelning.

- Man kan använda ett sådant diagram för att avgöra om mätningar verkar komma från en normalfördelning eller inte. Man kan även skatta väntevärde samt standardavvikelse för normalfördelad data.
- Sortera mätningarna från minsta till största. Sortera deras värden som x-axeln i diagrammet, x_i blir värdet på den mätning som var nummer i i ordningen från minst till störst.
- Skatta empiriska fördelningsfunktionen genom att ge varje mätning en sannolikhet beroende på dess plats i den sorterade ordningen, $p_i = \frac{i-0.5}{k}$.
- Är datan normalfördelad skall punkterna (x_i, p_i) nu ligga (approximativt) på en linje.

- Vi kan använda ett normalfördelningsdiagram för att avgöra vilka skattade effekter som är signifikanta.
- Först skattar vi alla effekter. Sedan sorterar vi dessa effekter och räknar ut motsvarande sannolikhetsvärde så att vi kan rita in dem i normalfördelningsdiagrammet.
- Vi tittar sedan på vilka punkter som ser ut att ligga på en linje och drar ett streck genom dessa. De punkter som inte ligger nära linjen anser vi är signifikanta effekter då de inte ser ut att följa referensfördelningen.

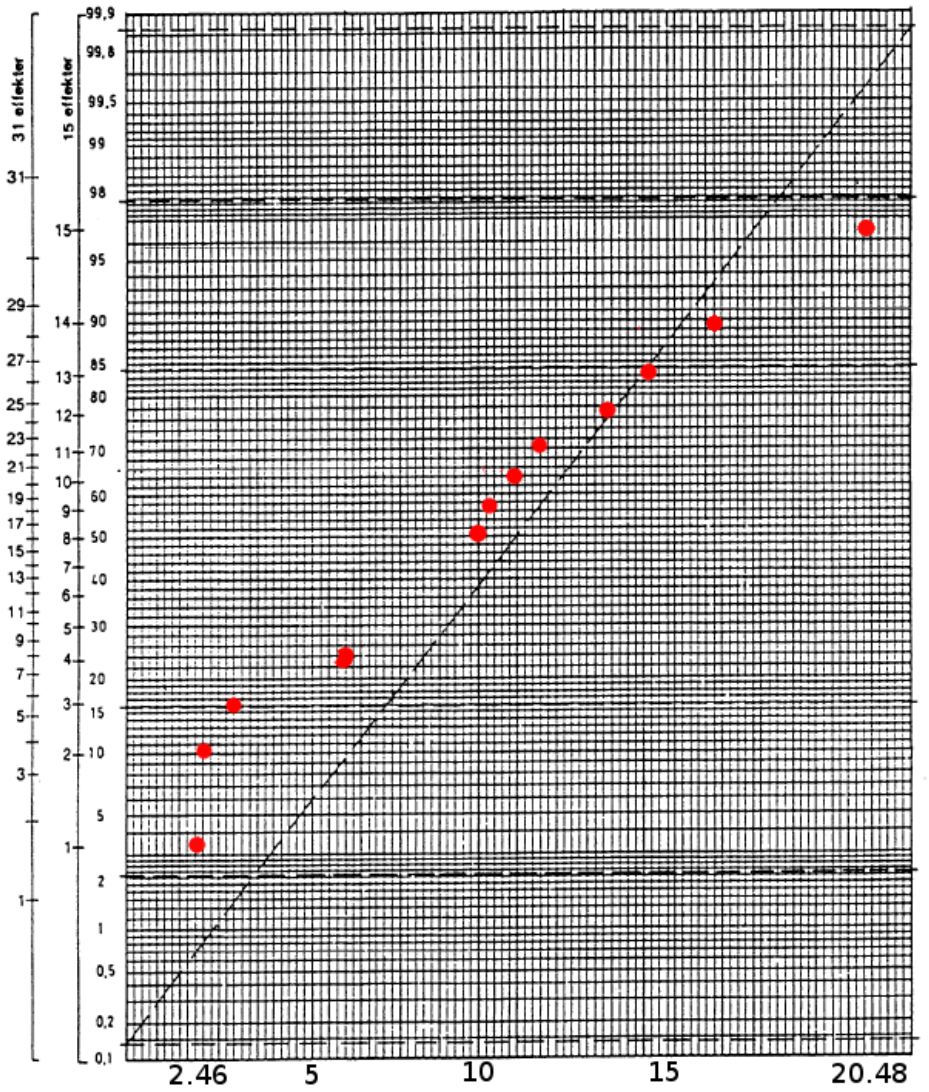
Exempel

i	x_i	p_i
1	2.46	0.033
2	2.52	0.100
3	3.46	0.167
4	5.99	0.233
5	7.45	0.3
6	9.02	0.367
7	9.14	0.433
8	9.98	0.500
9	10.18	0.567
10	11.21	0.633
11	12.34	0.70
12	13.44	0.767
13	14.23	0.833
14	15.98	0.90
15	20.48	0.967



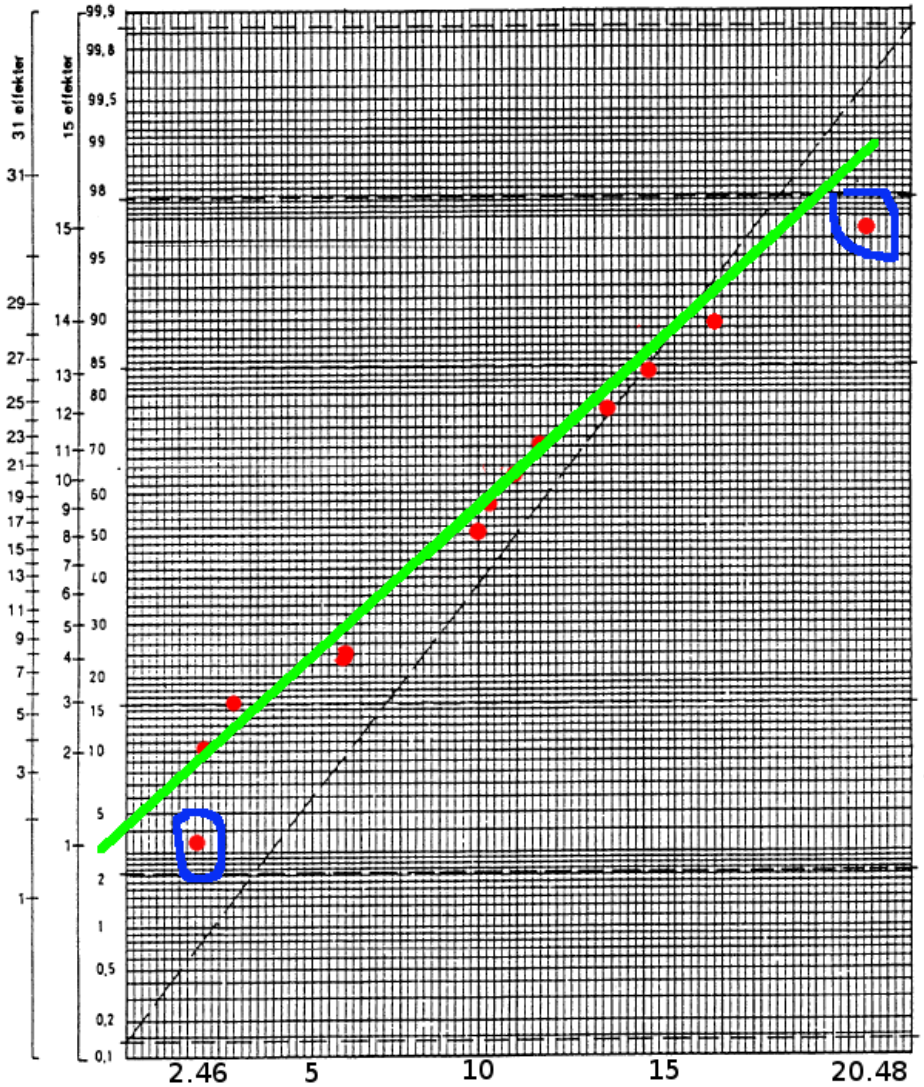
Exempel

i	x_i	p_i
1	2.46	0.033
2	2.52	0.100
3	3.46	0.167
4	5.99	0.233
5	7.45	0.3
6	9.02	0.367
7	9.14	0.433
8	9.98	0.500
9	10.18	0.567
10	11.21	0.633
11	12.34	0.70
12	13.44	0.767
13	14.23	0.833
14	15.98	0.90
15	20.48	0.967



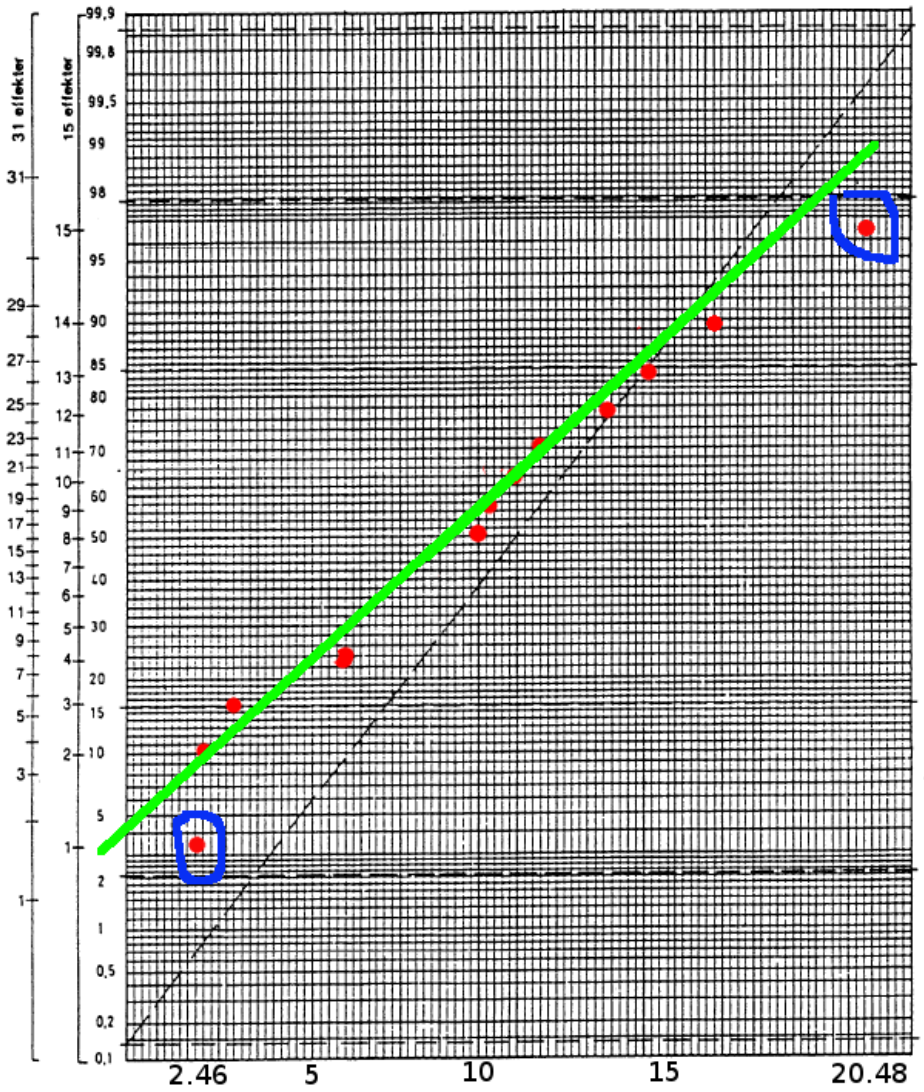
Exempel

i	x_i	p_i
1	2.46	0.033
2	2.52	0.100
3	3.46	0.167
4	5.99	0.233
5	7.45	0.3
6	9.02	0.367
7	9.14	0.433
8	9.98	0.500
9	10.18	0.567
10	11.21	0.633
11	12.34	0.70
12	13.44	0.767
13	14.23	0.833
14	15.98	0.90
15	20.48	0.967



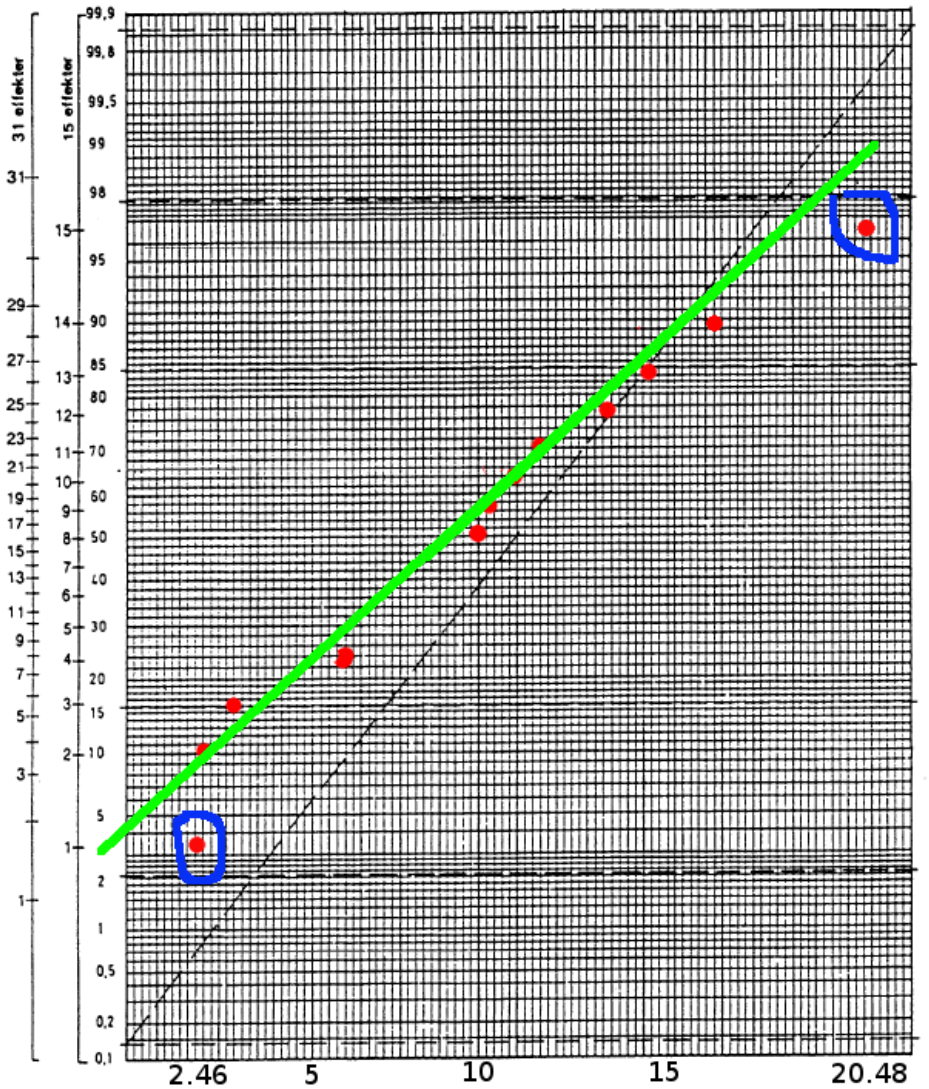
Exempel

Väntevärdet skattas genom att ta det x-värde på det gröna strecket som motsvarar y-värdet 0.5. I det här fallet motsvarar det ungefär 9.
Datan i det här exemplet är genererad slumpmässigt. För en riktig referensfördelning borde väntevärdet vara nära 0.



Exempel

Standardavvikelsen skattas genom att ta skillnaden mellan x-värdena där det gröna strecket korsar de tjocka streckade horisontella linjerna och dela med 4. $s \approx \frac{19+1}{4} = 5$.



Vi vet att referensfördelningen skall ha väntervärdet 0. Vi kan använda den kunskapen när vi ritar ut vårt streck i normalfördelningsdiagrammet.

Tips: anpassa streck för referensfördelningen

- Det anpassade strecket skall gå igenom punkten som har x-värdet 0 och y-värdet 50%.
- Strecket skall passa "så bra som möjligt" till de skattade effekterna omkring $x = 0$ då dessa är minst och därför troligast tillhör referensfördelningen.

- Fördelen med normalfördelningsdiagrammet är att vi inte behöver antaga vilka faktorer som är effektlösa och att vi inte heller kräver många olika mätningar inom samma grupp.
- Nackdelen är att tolkningen av normfördelningsdiagrammet är subjektiv (vad är tillräckligt långt från linjen och hur skall man välja linjen så den passar så bra som möjligt ihop med punkterna).

- Fördelen med normalfördelningsdiagrammet är att vi inte behöver antaga vilka faktorer som är effektlösa och att vi inte heller kräver många olika mätningar inom samma grupp.
- Nackdelen är att tolkningen av normfördelningsdiagrammet är subjektiv (vad är tillräckligt långt från linjen och hur skall man välja linjen så den passar så bra som möjligt ihop med punkterna).

En annan nackdel är att "normalfördelningsdiagrammet" är ett **väldigt** långt ord. Men inte lika långt som:
"nordvästersjökustartilleriflygspaningssimulatoranläggningsmateriel-underhållsuppföljningssystemdiskussionsinläggsförberedelsearbeten"
som tydligen skall vara svenskans längsta ord.

Sammanfattning

- Eftersom slumpen är inblandad så kommer inte skattade effekter vara exakt samma som riktiga effekter.
- Om vi känner till referensfördelningen så kan vi sätta en gräns för hur stor en skattad effekt måste vara för att vi skall tro att skattningen har någon signifikans.
- Vi antar normalfördelning hos referensfördelningen och antar att variansen är samma för alla mätgrupper.
- Variansen kan skattas på tre olika sätt.
  1. Skatta variansen i varje grupp och slå ihop dem.
  2. Antag att några effekter inte existerar och skatta variansen med hjälp av dessa.
  3. Rita ut de skattade effekterna på ett normalfördelningspapper och se vilka punkter som verkar ligga ungefär på ett streck.