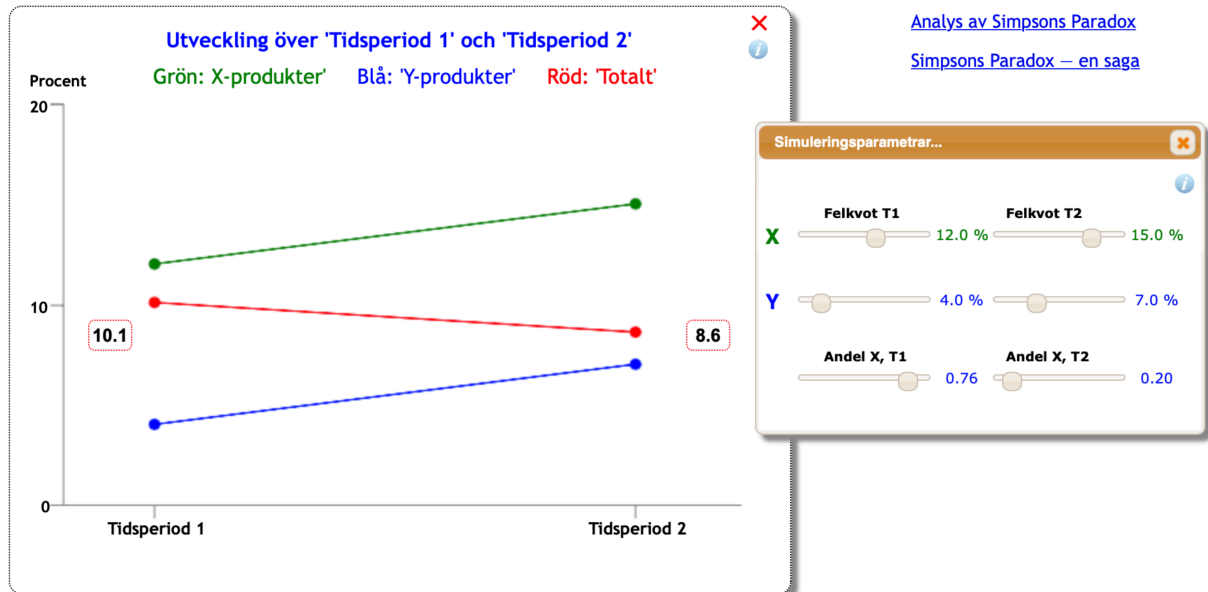


## Analys av Simpsons Paradox

Nedanstående diagram har skapats via länken [ovn.ing-stat.se/simpson/simpson2.php](http://ovn.ing-stat.se/simpson/simpson2.php). Diagrammet styrs av sliderna i parameterfönstret till höger.

Den gröna respektive blå linjen representerar en förändring över två tidsperioder. Den röda linjen visar totalt resultat och är en sammanvägning över respektive tidsperiod. Om 'sammanvägningensvikterna' förändras mellan T1 och T2 kan paradoxen, att den röda linjen visar annat resultat, uppstå (vikterna i nedanstående bild är 0.76, (1 - 0.76), 0.20 och (1 - 0.20).)

Se också formler under nedanstående bild.



## Beräkning

$p_{res}$	Resultat
$p_X$	Felkvot X-produkter
$p_Y$	Felkvot Y-produkter
$A_X$	Andel X-produkter
$A_Y = (1 - A_X)$	Andel Y-produkter

Tidsperiod 1

$$p_{res} = p_X \cdot A_X + p_Y \cdot (1 - A_X) = 12 \cdot 0.76 + 4 \cdot (1 - 0.76) \approx 10.1\%$$

Tidsperiod 2

$$p_{res} = p_X \cdot A_X + p_Y \cdot (1 - A_X) = 15 \cdot 0.20 + 7 \cdot (1 - 0.20) \approx 8.6\%$$

Dessa två värden (10.1% och 8.6%) visas också i grafen.

## Tabell över ansökningar till lärosäten

Datamängden visas i två olika perspektiv i nedanstående tabeller. Översta tabellen tycks visa att kvinnor verkar diskriminerade. Den undre tabellen visar dock exakt samma antagningsprocent för varje institution.

Den mellersta termen är lika i de två uttrycken nedan. Det framgår att många kvinnor (300) ansökte till institutionen med bara 25% antagning medan många män (300) ansökte till den institution som hade hög antagningsprocent. Detta gör att totalresultatet ser orättvis ut.

Tabell – ansökan till universitetet...

### Ansökningar till lärosäte

Kön	Antal sökande	Antal antagna	Kvot (%)
Män	600	350	58.3
Kvinnor	600	250	41.6

Nu finns det ingen könsskillnad och det är exakt samma antagningskvot till varje institution.  
(Se tabell 1B nedan)

Tabell 1A

Institution	Kön	Antal sökande	Antal antagna	Kvot (%)
I	Män	100	25	25
	Kvinnor	300	75	25
II	Män	200	100	50
	Kvinnor	200	100	50
III	Män	300	225	75
	Kvinnor	100	75	75

Tabell 1B

Män

$$p_{res} = 0.25 \cdot 100 / 600 + 0.50 \cdot 200 / 600 + 0.75 \cdot 300 / 600 \approx 58.3\%$$

Kvinnor

$$p_{res} = 0.25 \cdot 300 / 600 + 0.50 \cdot 200 / 600 + 0.75 \cdot 100 / 600 \approx 41.6\%$$

## Tabell över leveransförseningskvot

Datamängden visas i två olika perspektiv i nedanstående tabeller. Översta tabellen tycks visa att kvoten försämrats under tidsperiod 2, från 14.0% till 14.9%. Den undre tabellen visar dock en förbättring inom varje produktområde.

De två formlerna visar hur det totala resultatet i tabell 2A från uppdelningen i tabell 2B.

Tabell – felkvot mellan tidsperioder...

### Leveransförseningar

Tidsperiod	Antal order	Antal försenade	Kvot
1	880	123	0.140
2	1245	185	0.149

Nu visar varje produktslag en något ökad leveransprecision.  
(Se tabell 2B nedan)

Tabell 2A

Produkt	Tidsperiod 1			Tidsperiod 2		
	Antal order	Antal försenade	Kvot	Antal order	Antal försenade	Kvot
A	42	2	0.048	22	1	0.045
B	193	21	0.109	172	17	0.099
C	29	11	0.379	63	23	0.365
D	146	14	0.096	123	9	0.073
E	470	75	0.160	865	135	0.156
<b>Summa</b>	<b>880</b>	<b>123</b>	<b>0.140</b>	<b>1245</b>	<b>185</b>	<b>0.149</b>

Tabell 2B

*Tidsperiod 1*

$$p_{res} = 0.048 \cdot 42/880 + 0.109 \cdot 193/880 + 0.379 \cdot 29/880 + \dots \approx 14.0\%$$

*Tidsperiod 2*

$$p_{res} = 0.045 \cdot 42/880 + 0.099 \cdot 193/880 + 0.365 \cdot 29/880 + \dots \approx 14.9\%$$