

F5 - Introduktion till inferens

Statistisk översiktskurs

Ulf Högnäs

Statistiska institutionen
Stockholms universitet

April 1, 2025



Stockholm
University

- 1 Hypotesttest och the Randomization Test Procedure
- 2 Konfidensintervall och bootstrapping
- 3 Konfidensintervall och den centrala gränsvärdessatsen

Hypotesttest och the Randomization Test Procedure

Ett experiment från 70-talet

- Rosen, B., & Jerdee, T. H. (1974). Influence of sex role stereotypes on personnel decisions. *Journal of Applied Psychology*, 59(1), 9–14.

Ett experiment från 70-talet

- Rosen, B., & Jerdee, T. H. (1974). Influence of sex role stereotypes on personnel decisions. *Journal of Applied Psychology*, 59(1), 9–14.
- 48 bankchefer gick en managementutbildning

Ett experiment från 70-talet

- Rosen, B., & Jerdee, T. H. (1974). Influence of sex role stereotypes on personnel decisions. *Journal of Applied Psychology*, 59(1), 9–14.
- 48 bankchefer gick en managementutbildning
- De tilldelades var sin “personnel file” och bads avgöra om personen skulle befordras till mellanchefer

Ett experiment från 70-talet

- Rosen, B., & Jerdee, T. H. (1974). Influence of sex role stereotypes on personnel decisions. *Journal of Applied Psychology*, 59(1), 9–14.
- 48 bankchefer gick en managementutbildning
- De tilldelades var sin “personnel file” och bads avgöra om personen skulle befordras till mellanchefer
- Varje fil var identisk, förutom personens kön

Ett experiment från 70-talet

- Rosen, B., & Jerdee, T. H. (1974). Influence of sex role stereotypes on personnel decisions. *Journal of Applied Psychology*, 59(1), 9–14.
- 48 bankchefer gick en managementutbildning
- De tilldelades var sin “personnel file” och bads avgöra om personen skulle befordras till mellanchefer
- Varje fil var identisk, förutom personens kön
- 24 chefer fick bedömma en man, 24 en kvinna

Table 11.1: Summary results for the sex discrimination study.

sex	decision		Total
	promoted	not promoted	
male	21	3	24
female	14	10	24
Total	35	13	48

Är denna snedfördelning en ren slump?

sex_discrimination i R

Du kan lätt skapa tabellen från förra bilden i **R**.

```
# körs bara en gång
install.packages("openintro")
# ladda paketet som innehåller sexdistrimineringsdata
library(openintro)
# titta på de 6 första raderna i sex_discrimination
head(sex_discrimination)
# skapa en korstabell
table(sex_discrimination$sex, sex_discrimination$decision)
```

Experimentets Resultat

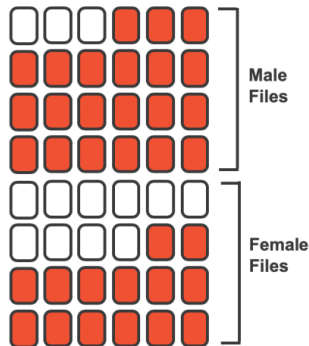
Andelen män som blev befördrade $\frac{21}{24} = 0.875$

Andelen kvinnor som blev befördrade $\frac{14}{24} \approx 0.583$

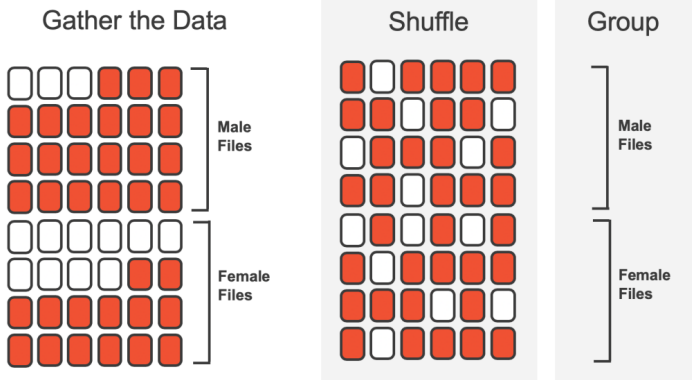
Skillnaden i andel $\frac{21}{24} - \frac{14}{24} = \frac{7}{24} \approx 0.292$

Simulering

- Vi tänker oss att ingen könsdiskriminering råder; 35 av 48 hade valt “promotion” för denna typ av kandidat, oavsett kön
- Vi lottar ut 35 promotions och ser hur könsfördelningen blir
- Hur ofta blir fördelningen mellan män och kvinnor 21 – 14, eller mer ojämn?

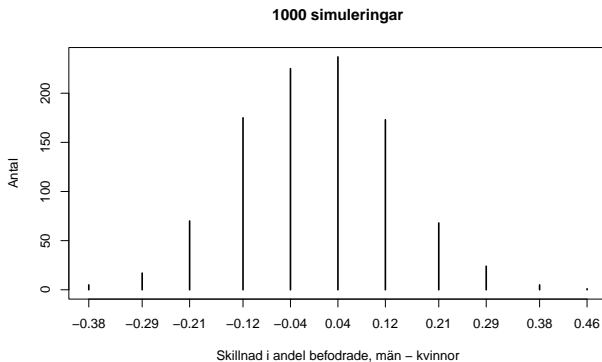


Simulering



Vi lottar. Denna gång blev det 18 – 17.

Exempel: Fördelning, 1000 Simuleringar



Skillnad	-0.38	-0.29	-0.21	-0.12	-0.04	0.04	0.12	0.21	0.29	0.38	0.46
Antal	5	17	70	175	225	237	173	68	24	5	1

Nollhypotesen och den alternativa hypotesen

Noll- och alternativa hypoteser

Nollhypotesen H_0 (eng. the null hypothesis) det skeptiska perspektivet; hypotesen att ingen skillnad från ett visst skeptiskt antagande existerar i verkligheten

Den alternativa hypotesen H_A (eng. the alternative hypothesis) ett alternativt påstående som försöket eller undersökningen eventuellt kan hitta stöd för

Nollhypotesen och den alternativa hypotesen

- Nollhypotesen H_0
 - Den genomsnittliga skillnaden skulle vara noll om försöket upprepades med andra chefer
 - Personens kön påverkar inte chansen att bli befördrad, om allt annat är lika

Nollhypotesen och den alternativa hypotesen

- Nollhypotesen H_0
 - Den genomsnittliga skillnaden skulle vara noll om försöket upprepades med andra chefer
 - Personens kön påverkar inte chansen att bli befordrad, om allt annat är lika
- Den alternativa hypotesen H_A
 - (the alternative hypothesis) - den genomsnittliga skillnaden skulle större än noll om försöket upprepades med andra chefer
 - Att vara kvinna påverkar kandidatens chans att bli befordrad negativt, om allt annat är lika

- Ett resultat som inträffar mer sällan än 5% av fallen givet antagandet att nollhypotesen är sann anses ofta vara minimum för att vi ska förkasta nollhypotesen

¹I boken står det cirka 2%. Varför skiljer det sig från vårt resultat?

- Ett resultat som inträffar mer sällan än 5% av fallen givet antagandet att nollhypotesen är sann anses ofta vara minimum för att vi ska förkasta nollhypotesen
- Gränsen beror på individer och sammanhang. Inom astrofysik är man 100,000 gånger strängare än så

¹I boken står det cirka 2%. Varför skiljer det sig från vårt resultat?

- Ett resultat som inträffar mer sällan än 5% av fallen givet antagandet att nollhypotesen är sann anses ofta vara minimum för att vi ska förkasta nollhypotesen
- Gränsen beror på individer och sammanhang. Inom astrofysik är man 100,000 gånger strängare än så
- I vårt fall: vi fick ett resultat som inträffar ungefär 3% av gångerna när nollhypotesen är sann¹

¹I boken står det cirka 2%. Varför skiljer det sig från vårt resultat?

- Ett resultat som inträffar mer sällan än 5% av fallen givet antagandet att nollhypotesen är sann anses ofta vara minimum för att vi ska förkasta nollhypotesen
- Gränsen beror på individer och sammanhang. Inom astrofysik är man 100,000 gånger strängare än så
- I vårt fall: vi fick ett resultat som inträffar ungefär 3% av gångerna när nollhypotesen är sann¹
- Skulle ni förkasta nollhypotesen i detta fall?

¹I boken står det cirka 2%. Varför skiljer det sig från vårt resultat?

Staistika

(eng. statistic or test statistic) Ett tal som sammanfattar datamaterialet

I vår undersökning använder vi oss av skillnaden i andel befordrade, män - kvinnor. Denna skillnad i andel är vår statistika.

$$\frac{21}{24} - \frac{14}{24} = \frac{7}{24} \approx 0.292$$

p-värde

Sannolikheten att observera data som är minst lika gynnsamma för alternativhypotesen som vårt nuvarande datamaterial, givet att nollhypotesen är sann^a

^aChat-GPTs översättning av bokens definition (s. 189)

p-värde

Sannolikheten att observera data som är minst lika gynnsamma för alternativhypotesen som vårt nuvarande datamaterial, givet att nollhypotesen är sann^a

^aChat-GPTs översättning av bokens definition (s. 189)

Vi uppskattar att om nollhypotesen är sann så skulle sannolikheten att se en skillnad på 0.292 eller mer vara 0.03 (3%). Detta är vårt p-värde.

Statistisk signifikans

Vi säger att vårt resultat är **statistiskt signifikant** (eng. statistically discernible) om p-värdet är under den gräns som vi själva valt

Statistisk signifikans

Vi säger att vårt resultat är **statistiskt signifikant** (eng. statistically discernible) om p-värdet är under den gräns som vi själva valt

Gränsen väljs innan datamaterialet tagits fram. Om vi har valt 5% som gräns så har vi hittat statistiskt signifikant stöd för den alternativa hypotesen *Att vara kvinna påverkar kandidatens chans att bli befordrad negativt, om allt annat är lika*

The Randomization Test Procedure

Från boken (s. 191)

- 1 Formulera hypoteser som stämmer med din forskningsfråga

The Randomization Test Procedure

Från boken (s. 191)

- 1 Formulera hypoteser som stämmer med din forskningsfråga
- 2 Samla in data (observationsstudie eller experiment)

The Randomization Test Procedure

Från boken (s. 191)

- 1 Formulera hypoteser som stämmer med din forskningsfråga
- 2 Samla in data (observationsstudie eller experiment)
- 3 Skapa ett slumpmässigt försök som bygger på att nollhypotesen stämmer

The Randomization Test Procedure

Från boken (s. 191)

- 1 Formulera hypoteser som stämmer med din forskningsfråga
- 2 Samla in data (observationsstudie eller experiment)
- 3 Skapa ett slumpmässigt försök som bygger på att nollhypotesen stämmer
- 4 Analysera data från detta försök; beräkna p-värdet givet din data

The Randomization Test Procedure

Från boken (s. 191)

- 1 Formulera hypoteser som stämmer med din forskningsfråga
- 2 Samla in data (observationsstudie eller experiment)
- 3 Skapa ett slumpmässigt försök som bygger på att nollhypotesen stämmer
- 4 Analysera data från detta försök; beräkna p-värdet givet din data
- 5 Dra en slutsats

Konfidensintervall och bootstrapping

Exempel: Politisk marknadsföring

- Blue Rose Research testade tusentals Kamala Harris-reklamsnuttar inför valet 2024
- Metod: 500 personer fick se en viss reklamsnutt, en kontrollgrupp på 500 personer såg inte reklamsnutt
- Samtliga 1000 deltagare fick svara på enkätfrågor
- Detta är ett exempel på en **randomized controlled trial** (sv. randomiserad kontrollerad studie)

In our data from last cycle, economic messaging that connected directly to voters lives tended to be most effective

Top testing ads most often:

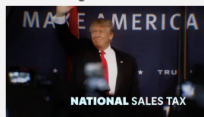
1. Acknowledging voters feeling squeezed by rising costs, showed how Harris would **improve voters' economic lives**.
2. Show that she has real plans that she would follow through on as President – vs. **how Trump was**

Direct-to-camera ad of VP Harris talking about lowering costs



VP: I get it. The cost of rent, groceries, and

Contrasting Trump raising prices and looking out for billionaires



VP: Building that middle class will be a defining

Ett problem med hypotestest

- Vi kan vara ganska säkra på att professionell politisk reklam har åtminstone någon positiv effekt (i genomsnitt)

Ett problem med hypotestest

- Vi kan vara ganska säkra på att professionell politisk reklam har åtminstone någon positiv effekt (i genomsnitt)
- Om vi genomför en tillräckligt stor undersökning kommer vi i så fall vi alltid att kunna förkasta nollhypotesen

Ett problem med hypotestest

- Vi kan vara ganska säkra på att professionell politisk reklam har åtminstone någon positiv effekt (i genomsnitt)
- Om vi genomför en tillräckligt stor undersökning kommer vi i så fall vi alltid att kunna förkasta nollhypotesen
- Frågan *hur effektiv är reklamsnutten?* är mer intressant

Hypothesetest vs. konfidensintervall

Hypotestest

Hypothesetest vs. konfidensintervall

Hypotestest

- Minskar ett visst vaccin risken att få malaria?

Hypothesetest vs. konfidensintervall

Hypotestest

- Minskar ett visst vaccin risken att få malaria?
- Har stödet för vänsterpartiet ökat sen förra mätningen?

Hypothestest vs. konfidensintervall

Hypotestest

- Minskar ett visst vaccin risken att få malaria?
- Har stödet för vänsterpartiet ökat sen förra mätningen?
- Är depression vanligare hos unga vuxna som använder appen TikTok än hos unga vuxna som inte använder TikTok?

Hypothestest vs. konfidensintervall

Hypotestest

- Minskar ett visst vaccin risken att få malaria?
- Har stödet för vänsterpartiet ökat sen förra mätningen?
- Är depression vanligare hos unga vuxna som använder appen TikTok än hos unga vuxna som inte använder TikTok?

Hypothestest vs. konfidsintervall

Hypotestest

- Minskar ett visst vaccin risken att få malaria?
- Har stödet för vänsterpartiet ökat sen förra mätningen?
- Är depression vanligare hos unga vuxna som använder appen TikTok än hos unga vuxna som inte använder TikTok?

Konfidsintervall

- Hur mycket minskar vaccinet risken att få malaria?

Hypothestest vs. konfidensintervall

Hypotestest

- Minskar ett visst vaccin risken att få malaria?
- Har stödet för vänsterpartiet ökat sen förra mätningen?
- Är depression vanligare hos unga vuxna som använder appen TikTok än hos unga vuxna som inte använder TikTok?

Konfidensintervall

- Hur mycket minskar vaccinet risken att få malaria?
- Hur mycket har stödet för vänsterpartiet ökat sen förra mätningen?

Hypothestest vs. konfidensintervall

Hypotestest

- Minskar ett visst vaccin risken att få malaria?
- Har stödet för vänsterpartiet ökat sen förra mätningen?
- Är depression vanligare hos unga vuxna som använder appen TikTok än hos unga vuxna som inte använder TikTok?

Konfidensintervall

- Hur mycket minskar vaccinet risken att få malaria?
- Hur mycket har stödet för vänsterpartiet ökat sen förra mätningen?
- Hur mycket vanligare är depression hos unga som använder appen TikTok än hos unga som inte använder TikTok?

Punktskattning

(eng. point estimate) Ett enskilda värde som används för att uppskatta värdet på en populationsparameter

Följande värden är punktskattningar om de används som en bästa gissning för värdet i en större population

Punktskattning

(eng. point estimate) Ett enskilda värde som används för att uppskatta värdet på en populationsparameter

Följande värden är punktskattningar om de används som en bästa gissning för värdet i en större population

- Andelen män som befordrades i stickprovet (0.875)

Punktskattning

(eng. point estimate) Ett enskilda värde som används för att uppskatta värdet på en populationsparameter

Följande värden är punktskattningar om de används som en bästa gissning för värdet i en större population

- Andelen män som befordrades i stickprovet (0.875)
- Skillnaden i andel, andel män minus andel kvinnor (0.293)

Punktskattning

(eng. point estimate) Ett enskilt värde som används för att uppskatta värdet på en populationsparameter

Följande värden är punktskattningar om de används som en bästa gissning för värdet i en större population

- Andelen män som befordrades i stickprovet (0.875)
- Skillnaden i andel, andel män minus andel kvinnor (0.293)
- Medelåldern i ett stickprov av studenter vid Södertörns högskola (t.ex. 24.1 år)

Statistika

Andelen män som beforderades i stickprovet

Skillnaden i andel, andel män minus andel kvinnor

Medelåldern i ett stickprov av studenter vid Södertörns högskola

Skattad populationsparameter

Sannolikheten att en slumpmässigt vald bankchef ska befordra en manlig anställd (som i experimentet)

Skillnaden i sannolikhet, män minus kvinnor

Medelåldern hos studenter vid Södertörns högskola

Konfidensintervall

(eng. confidence interval) ett intervall av värden som med en viss sannolikhet (t.ex. 95%) ska innehålla det sanna värdet på en populationsparameter

Konfidensintervall

(eng. confidence interval) ett intervall av värden som med en viss sannolikhet (t.ex. 95%) ska innehålla det sanna värdet på en populationsparameter

Exempel

- Sannolikheten att slumpmässigt vald bankchef ska befordra en manlig anställd som den i experimentet $(0.675, 0.967)$

Konfidensintervall

(eng. confidence interval) ett intervall av värden som med en viss sannolikhet (t.ex. 95%) ska innehålla det sanna värdet på en populationsparameter

Exempel

- Sannolikheten att slumpmässigt vald bankchef ska befordra en manlig anställd som den i experimentet $(0.675, 0.967)$
- Skillnaden i sannolikheter, män minus kvinnor $(0.0125, 0.571)$

Konfidensintervall

(eng. confidence interval) ett intervall av värden som med en viss sannolikhet (t.ex. 95%) ska innehålla det sanna värdet på en populationsparameter

Exempel

- Sannolikheten att slumpmässigt vald bankchef ska befordra en manlig anställd som den i experimentet (0.675, 0.967)
- Skillnaden i sannolikheter, män minus kvinnor (0.0125, 0.571)
- Medelåldren hos studenter vid S.H. (23.3, 24.9)

Medical Consultant Case Study (s. 198)

- 10% av patienter som genomgår levertransplantation drabbas av komplikationer
- En medicinsk konsult hävdar att hon är effektiv eftersom endast 3 av hennes 62 levertransplantationspatienter ($3/62 \approx 0.048 = 4.8\%$) drabbats av komplikationer

Medical Consultant Case Study (s. 198)

- 10% av patienter som genomgår levertransplantation drabbas av komplikationer
- En medicinsk konsult hävdar att hon är effektiv eftersom endast 3 av hennes 62 levertransplantationspatienter ($3/62 \approx 0.048 = 4.8\%$) drabbats av komplikationer
- 4.8% är en punktskattning av risken för konsultens framtida patienter

Medical Consultant Case Study (s. 198)

- 10% av patienter som genomgår levertransplantation drabbas av komplikationer
- En medicinsk konsult hävdar att hon är effektiv eftersom endast 3 av hennes 62 levertransplantationspatienter ($3/62 \approx 0.048 = 4.8\%$) drabbats av komplikationer
- 4.8% är en punktskattning av risken för konsultens framtida patienter
- Vi har bara sett 62 patienter från en oändlig population

Medical Consultant Case Study (s. 198)

- 10% av patienter som genomgår levertransplantation drabbas av komplikationer
- En medicinsk konsult hävdar att hon är effektiv eftersom endast 3 av hennes 62 levertransplantationspatienter ($3/62 \approx 0.048 = 4.8\%$) drabbats av komplikationer
- 4.8% är en punktskattning av risken för konsultens framtida patienter
- Vi har bara sett 62 patienter från en oändlig population
- Vi vill beräkna ett konfidensintervall för risken för komplikationer hos framtida patienter

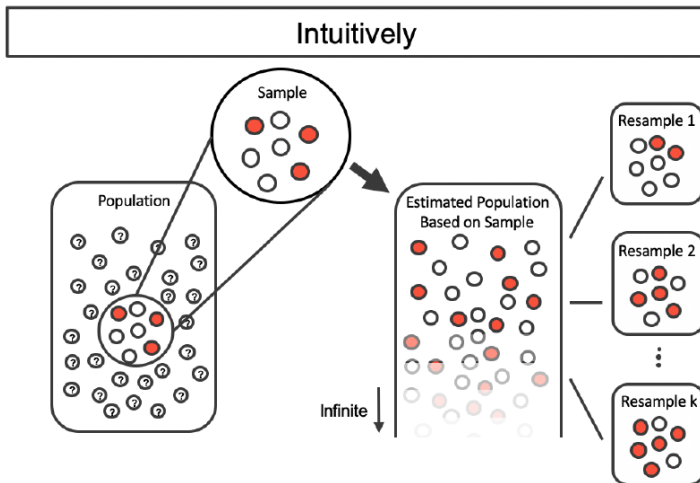


Figure: Från s. 200

Idé

- Vi har bara sett 62 fall

Idé

- Vi har bara sett 62 fall
- Vi skapar en oändlig population genom att kopiera dessa 62 fall oändligt många gånger

Idé

- Vi har bara sett 62 fall
- Vi skapar en oändlig population genom att kopiera dessa 62 fall oändligt många gånger
- Vi tar upprepade stickprov av storlek 62 från denna oändliga population

Idé

- Vi har bara sett 62 fall
- Vi skapar en oändlig population genom att kopiera dessa 62 fall oändligt många gånger
- Vi tar upprepade stickprov av storlek 62 från denna oändliga population
- För varje sådant stickprov beräknar vi andelen med komplikationer

Idé

- Vi har bara sett 62 fall
- Vi skapar en oändlig population genom att kopiera dessa 62 fall oändligt många gånger
- Vi tar upprepade stickprov av storlek 62 från denna oändliga population
- För varje sådant stickprov beräknar vi andelen med komplikationer
- Vi beräknar 2.5:e och 97.5:e percentilen bland alla stickprov

Idé

- Vi har bara sett 62 fall
- Vi skapar en oändlig population genom att kopiera dessa 62 fall oändligt många gånger
- Vi tar upprepade stickprov av storlek 62 från denna oändliga population
- För varje sådant stickprov beräknar vi andelen med komplikationer
- Vi beräknar 2.5:e och 97.5:e percentilen bland alla stickprov
- ... men istället för att göra kopior så gör vi slumpmässigt urval med återläggning

Bootstrap och percentiler

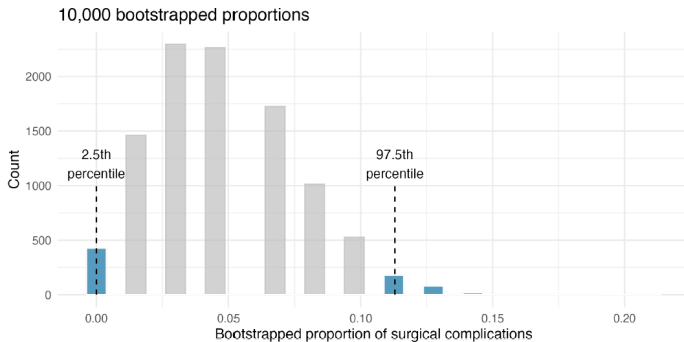


Figure: 95% av andelarna från vår bootstrapping ligger mellan 0% och 11.3%

Vad kan vi säga om konsultens påstående?

Vad kan vi säga om konsultens påstående?

- Vi kan inte säkert säga att den “sanna” risken för hennes framtida patienter är mindre än 10% i genomsnitt

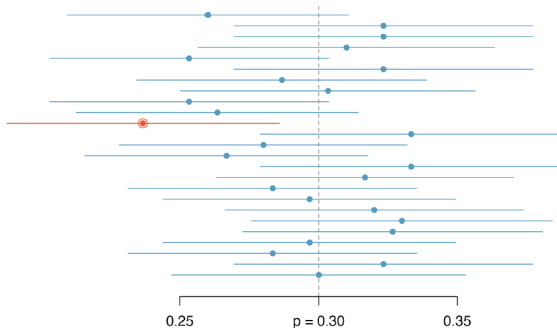
Vad kan vi säga om konsultens påstående?

- Vi kan inte säkert säga att den “sanna” risken för hennes framtida patienter är mindre än 10% i genomsnitt
- Det är osannolikt att risken är större än det amerikanska genomsnittet

Vad kan vi säga om konsultens påstående?

- Vi kan inte säkert säga att den “sanna” risken för hennes framtida patienter är mindre än 10% i genomsnitt
- Det är osannolikt att risken är större än det amerikanska genomsnittet
- Men som boken påpekar behöver förklaringen inte vara att konsulten är skicklig

Konfidensintervall och sannolikhet



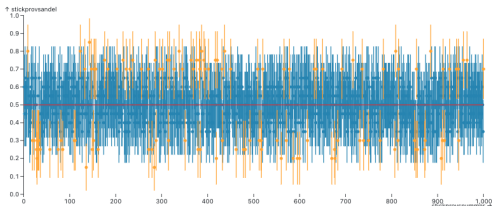
Ett 95%-igt konfidensintervall ska fånga den verkliga parametern 19 fall av 20. Varje linje motsvarar ett stickprov med tillhörande konfidensintervall

Konfidsensintervall för proportion, widget

Konfidsensintervall för en andel - normalapproximation



Av totalt 1000 stickprov innehöll 878 st (**87.800%**) av de 90%-iga konfidsensintervallen den sanna populationsandelen $p = 0.5$.



Prova Mattias Widget

Konfidensintervall och den centrala gränsvärdessatsen

Konfidensintervall och den centrala gränsvärdessatsen

Idé

Idé

- När stickprovsstorleken ökar närmar sig stickprovsmedelvärdets fördelning en normalfördelning

Idé

- När stickprovsstorleken ökar närmar sig stickprovsmedelvärdets fördelning en normalfördelning
- Stickprovsmedelvärdets varians kan skattas med stor noggrannhet m.h.a. stickprovet

Idé

- När stickprovsstorleken ökar närmar sig stickprovsmedelvärdets fördelning en normalfördelning
- Stickprovsmedelvärdets varians kan skattas med stor noggrannhet m.h.a. stickprovet
- Enligt 68-95-99.7-regeln hamnar Stickprovsmedelvärdet inom två standardavvikelser från populationsparametern med sannolikhet 95%

Konfidensintervall och den centrala gränsvärdessatsen

Idé

- När stickprovsstorleken ökar närmar sig stickprovsmedelvärdets fördelning en normalfördelning
- Stickprovsmedelvärdets varians kan skattas med stor noggrannhet m.h.a. stickprovet
- Enligt 68-95-99.7-regeln hamnar Stickprovsmedelvärdet inom två standardavvikelser från populationsparametern med sannolikhet 95%
- vi skapar ett intervall centrerat kring stickprovsmedelvärdet med samma bredd (se kludd på whiteboard)

Konfidensintervall och den centrala gränsvärdessatsen

Idé

- När stickprovsstorleken ökar närmar sig stickprovsmedelvärdets fördelning en normalfördelning
- Stickprovsmedelvärdets varians kan skattas med stor noggrannhet m.h.a. stickprovet
- Enligt 68-95-99.7-regeln hamnar Stickprovsmedelvärdet inom två standardavvikelser från populationsparametern med sannolikhet 95%
- vi skapar ett intervall centrerat kring stickprovsmedelvärdet med samma bredd (se kludd på whiteboard)
- detta intervall kommer att fånga populationsparametern med sannolikhet 95%

standardfel (SE)

Punktskattningar varierar från stickprov till stickprov, och denna variation kvantifieras med det som kallas **standardfel** (SE).

Standardfelet är detsamma som standardavvikelsen för den statistika vi undersöker. Det beskriver hur mycket ett punktskattning förväntas variera mellan olika stickprov.

²Följande fem bilder följer boken nära, från s. 225

standardfel (SE)

Punktskattningar varierar från stickprov till stickprov, och denna variation kvantifieras med det som kallas **standardfel** (SE).

Standardfelet är detsamma som standardavvikelsen för den statistika vi undersöker. Det beskriver hur mycket ett punktskattning förväntas variera mellan olika stickprov.

Standardfelet är i sig en skattning, beräknad från stickprovsdata. Hur vi bestämmer standardfelet beror på situationen, men oftast använder vi en formel som bygger på **centrala gränsvärdessatsen**.²

²Följande fem bilder följer boken nära, från s. 225

Felmarginal eller Margin of Error (ME)

Felmarginal

Avståndet $z^* \times SE$ kallas för **felmarginal**.

där z^* ett värde från normalfördelningen. Det vanligaste värdet är $z^* = 1.96$ (i 68-95-99.7-regeln är detta avrundat till 2). Detta ger en felmarginal som omfattar ungefär 95% av alla möjliga stickprovsskattningar

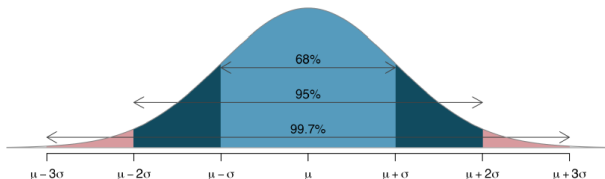


Figure: The 68-95-99.7 rule

Formel för ett 95%-igt konfidensintervall

Att konstruera ett 95%-konfidensintervall

När stickprovsfördelningen för en punktestimator (t.ex. \hat{p} eller \bar{x}) rimligen kan antas vara normalfördelad, kommer det observerade värdet att ligga inom 1.96 standardfel från det sanna värdet ungefär 95% av gångerna. Därför kan ett 95%-konfidensintervall konstrueras som

$$\text{punktskattning} \pm 1.96 \times SE$$

Vi kan då vara 95% säkra på att detta intervall kommer att fånga det sanna värdet i populationen.

Propotioner och “the success-failure contition”

the success-failure contition

Stickprovsfördelningen för \hat{p} (andel i stickprovet), baserad på ett stickprov av storlek n från en population med sann andel p , är ungefär normalfördelad när

- 1 Observationerna är oberoende, t.ex. från ett enkelt slumpmässigt urval.
- 2 Vi förväntar oss minst 10 lyckade och 10 misslyckade utfall: $np \geq 10$ och $n(1 - p) \geq 10$

I så fall är \hat{p} ungefär normalfördelad med

- Medelvärde p
- Standardfel $SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$

Formel för konfidsensintervall för ett proportion

Konfidsensintervall för en proportion, CGS

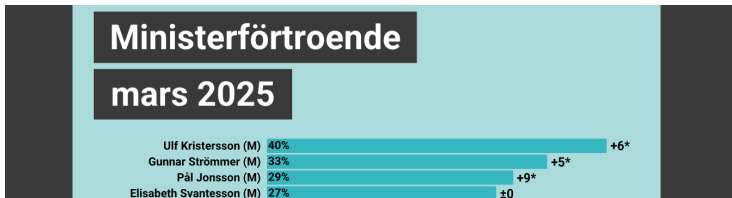
Om “the success-failure condition” är uppfyllt så kan det 95%-iga konfidsensintervallet för en proportion beräknas som

$$\hat{p} \pm 1.96 \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$$

där \hat{p} är stickprovsproportionen och n är stickprovsstorleken

Ett räkneexempel

- Novus undersökning omfattade 509 personer³
- “Vilket förtroende har du för följande ministrar?”
- 204 svarande gav Ulf Kristersson “ganska stort” eller “mycket stort”
- Beräkna ett 95%-igt konfidensintervall för andelen bland alla väljare som har ganska stort eller stort förtroende för Kristersson



³ egentligen 1005 personer, men alla fick inte frågan om Kristersson

① Beräkna \hat{p}

$$\hat{p} = \frac{204}{509} \approx 0.4008$$

- ① Beräkna \hat{p}

$$\hat{p} = \frac{204}{509} \approx 0.4008$$

- ② kontrollera “the success-failure condition”

$$n \cdot \hat{p} = 509 \cdot \frac{204}{509} = 204 \geq 10$$

$$n \cdot (1 - \hat{p}) = 509 \cdot \left(1 - \frac{204}{509}\right) = 305 \geq 10$$

OK. De förväntade antalen stort-ganska-stort / inte stort-ganska-stort i stickprovet är båda större än 10. Vi kan använda den centrala gränsvärdessatsen

- ③ Använd formeln på bild 33

$$\begin{aligned} & \hat{p} \pm 1.96 \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} \\ & 0.4008 \pm 1.96 \sqrt{\frac{0.4008(1 - 0.4008)}{509}} \\ & 0.4008 \pm 0.04257 \end{aligned}$$

- ③ Använd formeln på bild 33

$$\begin{aligned} & \hat{p} \pm 1.96 \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} \\ & 0.4008 \pm 1.96 \sqrt{\frac{0.4008(1 - 0.4008)}{509}} \\ & 0.4008 \pm 0.04257 \end{aligned}$$

- ④ Beräkna den lägre och den övre gränsen, avrunda

$$(0.36, 0.44)$$

Vi kan säga med 95 procents konfidens att mellan 36% och 44% av väljarna hade stort eller ganska stort förtroende för Kristersson, mars 2025

Mer om inferens

- Konfidensintervall
- Hypotestest
- p-värde