Resultatvärdering

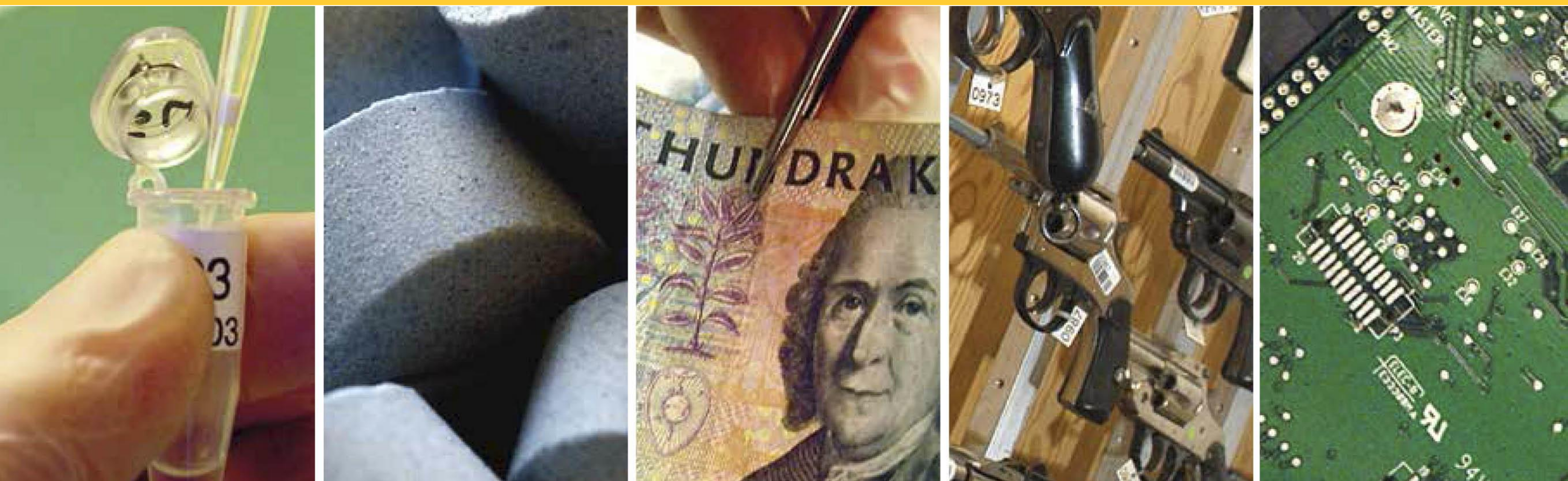
Polisen

Anders Nordgaard

Docent och Forensisk specialist, Nationellt Forensiskt Centrum

Adj. Universitetslektor i Statistik, Linköpings Universitet

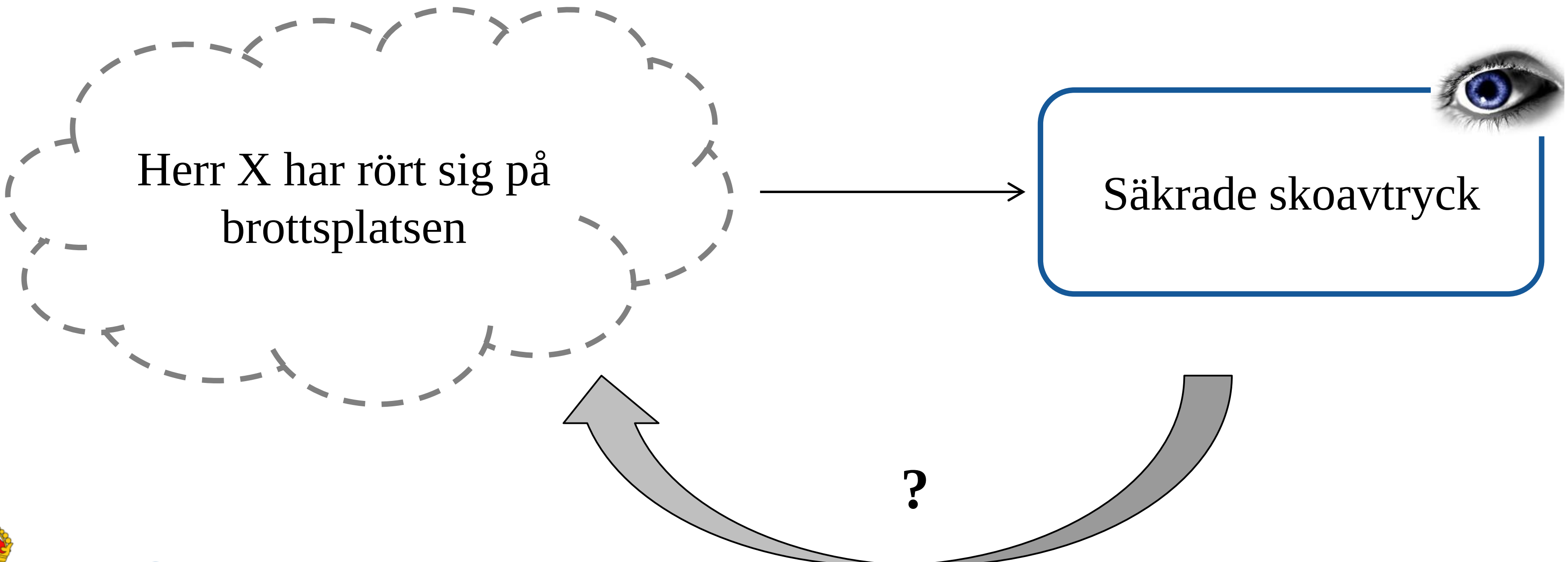
Doktorandkurs i IT-forensik, ISY våren 2018



Forensisk (*teknisk*) bevisning…

… är en *induktiv* slutledningsprocess.

> ”Dra slutsatser om vad som har hänt utifrån observerade konsekvenser av det som har hänt – det vi i efterhand ser”



Den grundläggande logiska teorin

Påståenden uppstår om vad som har hänt, vad som är sant.

Hypoteser

- Förövaren tog sig in i huset via vardagsrumsfönstret
- Blodet på golvet i köket kommer från den döde, som påträffades i sovrummet
- Overallen har varit i kontakt med bilsätet
- Datorn har använts för fildelning

Prövning av hypoteser görs två och två i taget.

Huvudhypotes – ***Alternativhypotes***

$\boldsymbol{H_h}$: Overallen har varit i kontakt med bilsätet

$\boldsymbol{H_a}$: Overallen har inte varit i kontakt med bilsätet

Polisen

När en hypotes uppstår finns en förhandsuppfattning om hur sannolikt det är att den hypotesen är sann

$P(\boldsymbol{H_h})$ ett tal mellan 0 och 1 (0 % och 100 %)

I de flesta fall innebär detta att sannolikheten för att alternativhypotesen är sann är

$$P(\boldsymbol{H_a}) = 1 - P(\boldsymbol{H_h})$$

Detta kan uttryckas i så kallade *förodds* för $\boldsymbol{H_h}$:

$$\frac{P(H_h)}{P(H_a)}$$

”Hur mycket som initialt talar för H_h i förhållande till hur mycket som initialt talar för H_a”

Föroddsen finns hos den som ”äger” frågeställningen!

Det forensiska laboratoriet äger aldrig frågeställningen!

Polisen

Bevisvärderingens/resultatvärderingens kärna:

Använd resultaten ***E*** från den forensiska undersökningen för att uppdatera sannolikheterna för $\boldsymbol{H_h}$ respektive $\boldsymbol{H_a}$ – uppdatera oddsen.

Bayes' sats (*regel, teorem*)

Resultatvärdet – hur mycket mer (eller mindre) sannolika resultaten är om $\boldsymbol{H_h}$ är sann jämfört med om $\boldsymbol{H_a}$ är sann

$$\frac{P(H_h|E)}{P(H_a|E)} = \frac{P(E|H_h)}{P(E|H_a)} \times \frac{P(H_h)}{P(H_a)}$$

Uppdaterade odds (*efterodds*) när resultaten ***E*** tagits i beaktande

Förodds innan resultaten ***E*** tagits i beaktande

Bevisvärderingens/resultatvärderingens kärna:

Använd resultaten ***E*** från den forensiska undersökningen för att uppdat

oddsen.

$P(H_h|E)$ står för sannolikheten att H_h är sann givet att resultaten E har observerats.

$P(E|H_h)$ står för sannolikheten att observera E givet att H_h är sann

Bayes'sa

r sannolika

är sann jämfört

$$\frac{P(H_h|E)}{P(H_a|E)} = \frac{P(E|H_h)}{P(E|H_a)} \times \frac{P(H_h)}{P(H_a)}$$

Uppdaterade odds (*efterodds*) när resultaten ***E*** tagits i beaktning

Förodds innan resultaten ***E*** tagits i beaktning

Den forensiska processens två faser

- Utredande fas (investigative mode)
- Värderande fas (evaluative mode)

Den utredande fasen har som mål en sammantagen kriminalteknisk bedömning av vad som hänt i ett brott (på brottsplatsen/erna, fyndplatsen/erna, etc.)

Specifika frågeställningar om undersökta spår och deras ev. kopplingar till misstänkta personer och/eller beslagtaget gods hanteras i den värderande fasen
- *kan bli avgörande för skuldfrågan*

…men faserna förekommer omväxlande under ärendets gång, dvs. det är inte den ena först och den andra därefter.

Den värderande fasen på laboratoriet

I utredningen uppstår specifika frågor rörande spår säkrade på brottsplats och personer/gods misstänkta att ha orsakat spåren. Det kan också handla om frågeställningar runt länkar mellan olika material.

Som regel måste dessa frågor hanteras i en laboratorieundersökning

- Var det den här skon som avsatte det säkrade skoavtrycket?
- Kommer blodet på golvet från den döde som påträffats på platsen?
- Var det den misstänkte som krossade rutan?
- Har det eldats med miljöfarligt bränsle på platsen?
- Kommer de två beslagtagna påsarna med amfetamin från samma tillverkningssats?
- …

Gemensamt för dessa typer av frågor är att de har de <u>ultimata</u> svaren Ja och Nej.

Svaren på sådana frågor förväntas kunna

– reda ut förhållanden på en brotts/fyndplats som avgör vilken riktning den fortsatta undersökningen ska ta

… om slutsatsen tolkas som Ja eller Nej av brottsplatsutredaren.

– användas som stöd för en hypotes om ett visst förlopp på brotts/fyndplatsen

– självständigt utgöra den tekniska bevisning som kopplar ett omstritt material till en person eller till ett annat material, eller till en viss klass

Svarar laboratoriet då alltid endera Ja eller Nej?

Polisen

Frågeställningar (och värdering) på tre olika nivåer

Källnivå:

Frågeställningen gäller om ett omstritt material

- härrör från en viss utpekad källa (*blod från en person?*, *glas från en fönsterruta?*)
- har samma ursprung som ett jämförelsematerial (*två beslag av amfetamin från samma tillverkningssats?*, *två skoavtryck avsatta med samma sko?*)
- tillhör en viss klass med avseende på sina egenskaper (*brännbar vätska?*, *falskt pass?*, *giftigt material?*, *heroin i pulvret?*)

Aktivitetsnivå:

Frågeställningen gäller om en viss aktivitet har pågått (som gett upphov till spåren). (*Har den misstänkte krossat glasrutan?*, *Har den misstänkte haft samlag med målsägaren?*, *Har den misstänkte suttit på bilsätet?*)

Brottsnivå:

Frågeställningen gäller om den tilltalade är skyldig till det brott hen åtalats för. Hanteras inte i den forensiska processen.

Hur ska det gå till i praktiken?

NFC får in en **Beställning av forensisk undersökning**

Beställning av forensisk undersökning 1 (2)

Uppdragsgivare
Polismyndigheten
BOX 123
123 11 SKURKSUND

NFC:s dnr

Undersökningsmaterial

Beslagsnummer	Beteckning	Materialbeskrivning	Begärda undersökn.	Jämför med material	Referensnummer
2015-5000-BG12492-2	A1/5000-15/G002	Kalsong	K12		5000-K12345-15
	A1/5000-15/G062	Skor	K12,S32		5000-K12345-15
2015-5000-BG12441-1	A1/5000-15/G071	Skor	K12,S32		5000-K12345-15
2015-5000-BG12760-1	A1/5000-15/G072	Kavaj	K10,K11,K12, M31		5000-K12345-15
2015-5000-BG12760-2	A1/5000-15/G073	Klänning	K10,K11,K12, M31		5000-K12345-15
2015-5000-BG12764-7	A1/5000-15/G085	Kniv med slida	K10,K12,S21, F50,M31		5000-K12345-15
	A1/5000-15/G087	6 st. stenar	K12, F50		5000-K12345-15
2015-5000-BG12798-1	A1/5000-15/G098	Jeans	K10,K12,K71, F50		5000-K12345-15
	A1/5000-15/S022	Fibrer säkrade med tejp - från vänster underben samt höger hands utsida	F50		5000-K12345-15
	A1/5000-15/S049	Miljöprov - anträffandeplats offret	K71		5000-K12345-15
	A1/5000-15/S059	Miljöprov på tejp, brottsplats, förvaring- och fyndplats	F50		5000-K12345-15
	A1/5000-15/S060	Skospår	S32		5000-K12345-15

Materialet önskas åter? Ja

Vi använder i denna föreläsning två av materialen (och en undersökning).

Undersökningar begärda och motsvarande material

PRODUKTKATALOG

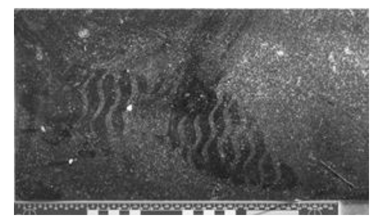
Spårundersökningar

Kod	Klartext	Benämning	Beskrivning
S00	Övrig spårundersökning	Övriga spårundersökningar, t.ex. fotavtryck (ej skoavtryck), handskavtryck och däckavtryck.	Om det inte finns någon kod som passar in på den efterfrågade undersökningen sätts koden till S00, som kompletteras med en skriftlig precisering av vilken typ av undersökning som önskas.
S11	Platsundersökning, spår	Platsundersökning, spår.	Bistå vid platsundersökning för hantering av skoavtryck, fotavtryck, verktygsspår o.s.v.
S21	Fingeravtrycksframkallning	Fingeravtrycksundersökning.	Fingeravtrycksframkallning på insända föremål. Jämförelse och identifiering av säkrade avtryck.
S22	Platsundersökning, fingeravtryck	Fingeravtrycksundersökning på brottsplats.	Bistå vid platsundersökning i samband med fingeravtrycksundersökningar
S31	Säkring/fotografering av skoavtryck	Säkring/fotografering av skoavtryck	Framkallning/förstärkning av skoavtryck med bl. a. fototekniska metoder.
S32	Skoavtryck	Skoavtrycksundersökning.	Jämförelse av säkrade spår med beslagtagna skor och/eller ev. andra säkrade spår. Framkallning/förstärkning av skospår ingår.
S41	Verktygsspår	Verktygsspårsundersökning	Jämförelse av säkrade spår med beslagtagna verktyg och/eller med andra säkrade spår. Jämförelse av präglingen i tabletter.
S42	Lås och nycklar	Lås och nycklar	Undersökning av lås och nycklar avseende dyrkning, manipulation och kopiering.

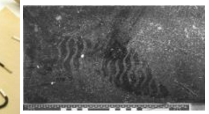
2015-5000-BG12492-2

A1/5000-15/G062 Skor

2015-5000-BG12798-1



A1/5000-15/S060 Skospår



- Vem är det som ”står bakom” beställningen?
- Vad vill ”beställaren” veta?

Ur NFC:s synvinkel är det förundersökningsledaren

S32	Skoavtryck	Skoavtrycksundersökning.	Jämförelse av säkrade spår med beslagtagna skor och/eller ev. andra säkrade spår. Framkallning/förstärkning av skospår ingår.

Den mest naturliga frågeställningen torde vara:

Har någon av dessa skor avsatt det säkrade skoavtrycket?

Dock kanske det är uppenbart att bara en av dem kan ha gjort det.

I förundersökningen borde man ha svar på följande:

Varför beslagtogs de här skorna?

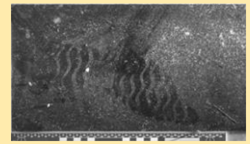
Hur mycket är det som inledningsvis talar för att det var de här skorna som avsatte avtrycket? Finns det något som talar emot detta?

Förundersökningen äger frågeställningen, inte laboratoriet!

Polisen

Resultatvärdering vid källnivåjämförelser

Ett säkrat skoavtryck och ett par beslagtagna skor.

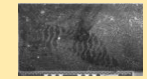


Resultat:

- Sulmönstret i skoavtrycket överensstämmer med sulmönstret hos vänsterskon
- En förslitning på vänsterskons sula överensstämmer i stort med motsvarande del av skoavtrycket

Vad betyder detta?

Ett säkrat skoavtryck och
ett par beslagtagna skor.



Har någon av dessa skor avsatt det säkrade skoavtrycket?

Vilken slags svar förväntar sig förundersökningsledaren (idealt)?

”Det var en av dessa skor som avsatte avtrycket” (Ja)

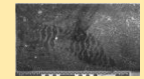
eller

”Det var ingen av dessa skor som avsatte avtrycket” (Nej)

Går det att svara Ja?

Polisen

Ett säkrat skoavtryck och ett par beslagtagna skor.



”Vi” kan svara Ja om…

…vi är 100 % säkra på att ingen annan sko kan ha åstadkommit detta avtryck

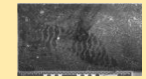
Hur ofta kan vi vara det?

”Vi” svarar Nej om…

…vi är 100 % säkra på att ingen av de beslagtagna skorna kan ha åstadkommit detta avtryck

Bägge dessa kategoriska svar kommer av att vi kan göra en säker uteslutning – övrig bevisning i frågeställningen får då ingen inverkan.

Ett säkrat skoavtryck och ett par beslagtagna skor.



Om vi nu inte kan göra en säker uteslutning…

Omforma den initiala frågeställningen…

Har någon av dessa skor avsatt det säkrade skoavtrycket?

…med hänsyn tagen till att endast en av skorna – vänsterskon kan ha gett avtrycket…

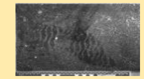
…till ett påstående, en ***huvudhypotes***

$\boldsymbol{H_h}$ **:** Beslagtagen sko G har avsatt det säkrade skoavtrycket S

Bevisvärdet hos hypotesen utgörs av <u>tilltron</u> till <u>hypotesen</u> – och dess relevans för den aktuella gärningsbeskrivningen.

Polisen

Ett säkrat skoavtryck och ett par beslagtagna skor.



Vi fokuserar på tilltron…

När tilltron inte är 100 % eller 0 % (uteslutningsfallen) måste man använda sannolikhetskalkyl.

Kan man forensiskt direkt bedöma sannolikheten för att hypotesen

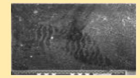
> $\boldsymbol{H_h}$ **:** Beslagtagen sko G har avsatt det säkrade skoavtrycket S

är sann?

> <u>Svar: Nej.</u>
> Sannolikheten bedöms utifrån en kombination av den forensiska bedömningen och annan information i utredningen – annan bevisning än den forensiska.

Polisen

Ett säkrat skoavtryck och ett par beslagtagna skor.



Annan information i utredningen

Forensisk bedömning (*vid laboratoriet*)

$\boldsymbol{H_h}$ **:** Beslagtagen sko G har avsatt det säkrade skoavtrycket S

Slutlig sannolikhet att $\boldsymbol{H_h}$ är sann

***H_h* :** Beslagtagen sko G har avsatt det säkrade skoavtrycket S

Annan information i utredningen

Gemensam bedömningsplattform

Forensisk bedömning (*vid laboratoriet*)

Alternativhypotes

t.ex. ***H_a* :** En annan sko har avsatt spåret

Väljs så att den omfattar alla relevanta alternativ till huvudhypotesen.

H_h **:** Beslagtagen sko G har avsatt det säkrade skoavtrycket S

H_a **:** En annan sko har avsatt spåret

Forensisk bedömning

Hur förväntat/sannolikt är…

- Sulmönstret i skoavtrycket överensstämmer med sulmönstret hos vänsterskon
- En förslitning på vänsterskons sula överensstämmer i stort med motsvarande del av skoavtrycket

…om ***huvudhypotesen*** är sann? $\Rightarrow P(\text{Resultaten} \mid H_h)$

…om ***alternativhypotesen*** är sann? $\Rightarrow P(\text{Resultaten} \mid H_a)$

$$\text{Resultatvärdet} = V = \frac{P(\text{Resultaten} \mid H_h)}{P(\text{Resultaten} \mid H_a)}$$

Polisen

$\boldsymbol{H_h}$ **:** Beslagtagen sko G har avsatt det säkrade skoavtrycket S

$\boldsymbol{H_a}$ **:** En annan sko har avsatt spåret

Om sannolikheter…

Sannolikheten för ***något*** är

graden av tilltro till detta ***något*** hos den som har att förhålla sig till detta ***något*** utifrån

den ***bakgrundsinformation*** som denne har och de ***antaganden***, som denne ska göra

– ett tal mellan 0 % och 100 % (mellan 0 och 1)

En sannolikheten betecknas med *P* (efter engelskans *p*robability)

P(***något*** | ***antaganden***, (*bakgrundsinformation*) **)**

$P(\text{Resultaten} \mid \boldsymbol{H_h})$

under antagandet att ***huvudhypotesen*** är sann

$P(\text{Resultaten} \mid \boldsymbol{H_a})$

under antagandet att ***alternativhypotesen*** är sann

Polisen

H_h **:** Beslagtagen sko G har avsatt det säkrade skoavtrycket S

H_a **:** En annan sko har avsatt spåret

Forensisk bedömning

$$\text{Resultatvärdet} = V = \frac{P\left(\text{Resultaten} \mid H_h\right)}{P\left(\text{Resultaten} \mid H_a\right)}$$

…anger

<u>om V är större än 1</u>

hur mycket ***mer*** sannolika Resultaten är om ***huvudhypotesen*** är sann jämfört med om ***alternativhypotesen*** är sann.

<u>om V är mindre 1</u>

hur mycket ***mindre*** sannolika Resultaten är om ***huvudhypotesen*** är sann jämfört med om ***alternativhypotesen*** är sann.

<u>om V är lika med 1</u>

att Resultaten är lika sannolika om ***huvudhypotesen*** är sann som om ***alternativhypotesen*** är sann.

Praktisk beräkning/bedömning av resultatvärdets storlek

För några forensiska ämnesområden finns omfattande relevanta bakgrundsdata och validerade matematiska stödmodeller för att ”beräkna” resultatvärdet (*t.ex. DNA- och glasjämförelser)*

I flertalet forensiska ämnesområden är dock ännu omfånget av bakgrundsdata och stödmodeller begränsat.

Ett forensiskt laboratorium bör dock ha ett enhetligt sätt att rapportera resultatvärden.

För forensiska undersökningar där stödmodeller inte har implementerats måste erfarenhets- och sakkunskapsbaserad (subjektiv) bedömning av resultatvärdets komponenter göras
$\Rightarrow$ Relativt grova uppskattningar av resultatvärdets storlek

$\Rightarrow$ All redovisning av resultatvärden från NFC - med eller utan användande av modell - görs på en gemensam graderad slutsatsskala!

NFC:s utlåtandeskala

Grad +4 Resultaten talar extremt starkt för att ...
Det bedöms vara extremt mycket mer sannolikt att få dessa resultat om huvudhypotesen är sann än om den alternativa hypotesen är sann.

Grad +3 Resultaten talar starkt för att ...
Det bedöms vara mycket mer sannolikt att få dessa resultat om huvudhypotesen är sann än om den alternativa hypotesen är sann.

Grad +2 Resultaten talar för att ...
Det bedöms vara mer sannolikt att få dessa resultat om huvudhypotesen är sann än om den alternativa hypotesen är sann.

Grad +1 Resultaten talar i någon mån för att ...
Det bedöms vara något mer sannolikt att få dessa resultat om huvudhypotesen är sann än om den alternativa hypotesen är sann.

Grad 0 Resultaten talar varken för eller emot att ...
Det bedöms vara ungefär lika sannolikt att få de erhållna resultaten om huvudhypotesen är sann som om den alternativa hypotesen är sann.

Grad -1 Resultaten talar i någon mån för att ... inte ...
Det bedöms vara något mer sannolikt att få dessa resultat om den alternativa hypotesen är sann än om huvudhypotesen är sann.

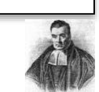
Grad -2 Resultaten talar för att ... inte ...
Det bedöms vara mer sannolikt att få dessa resultat om den alternativa hypotesen är sann än om huvudhypotesen är sann.

Grad -3 Resultaten talar starkt för att ... inte ...
Det bedöms vara mycket mer sannolikt att få dessa resultat om den alternativa hypotesen är sann än om huvudhypotesen är sann.

Grad -4 Resultaten talar extremt starkt för att ... inte ...
Det bedöms vara extremt mycket mer sannolikt att få dessa resultat om den alternativa hypotesen är sann än om huvudhypotesen är sann.

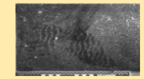
Gradernas koppling till resultatvärdets storlek:

Slutsats-grad	Resultatvärdets storlek	”Förklaring”
		Resultaten är…
+4	minst en miljon	…minst en miljon ggr sannolikare…
+3	mellan 6000 och en miljon	…minst 6000 ggr sannolikare…
+2	mellan 100 och 6000	…minst 100 ggr sannolikare…
+1	mellan 6 och 100	…minst 6 ggr sannolikare…
0	mellan 1/6 och 6	… ungefär lika sannolika…
		…om huvudhypotesen är sann jämfört med om den alternativa hypotesen är sann
–1	mellan 1/100 och 1/6	…minst 6 ggr sannolikare…
–2	mellan 1/6000 och 1/100	…minst 100 ggr sannolikare…
–3	mellan 1/(en miljon) och 1/6000	…minst 6000 ggr sannolikare…
–4	högst 1/(en miljon)	…minst en miljon ggr sannolikare…
		…om den alternativa hypotesen är sann jämfört med om huvudhypotesen är sann



Hur svarar NFC på frågeställningen?

Ett säkrat skoavtryck och ett par beslagtagna skor.



Ur sakkunnigutlåtandet:

Ändamål

Skoavtryck

Ändamålet är att undersöka om skoavtrycket A1/5000-15/S060, säkrat med avgjutning, har avsatts med skorna A1/5000-15/G062.

Undersökning och slutsats

A1/5000-15/S060	Skoavtryck avsatt på betongplatta *SKL:s materialnr: 201500001405*
Skoavtryck	**Resultat** Skoavtrycket A1/5000-15/S060 jämfördes med sulorna på skorna A1/5000-15/G062 och avtryck avsatta med dessa. Mönstertypen i det omstridda avtrycket överensstämde med mönstertypen i sulan på vänsterskon A1/5000-15/G062. Förslitningen i främre delen av sulan på vänsterskon överensstämde i stort med motsvarande del i det omstridda avtrycket. Inga särskiljande olikheter iakttogs. : ko. Det bedöms fullt rimligt att det omstridda avtrycket ska ha de aktuella egenskaperna om det har avsatts med vänsterskon. Avtrycket har en tydligt tecknad mönstertyp som tillhör de mindre vanligt förekommande mönstertyperna. Vidare är förslitningens utseende av sådant slag att den inte bedöms vara vanlig för den här typen av skosula. Andelen skor som skulle kunna avsätta ett sådant avtryck bedöms därför vara liten. **Slutsats** Resultaten talar för att skoavtrycket A1/5000-15/S060 har avsatts med vänsterskon A1/5000-15/G062 (Grad +2).

Slutsats

Resultaten talar för att skoavtrycket A1/5000-15/S060 har avsatts med vänsterskon A1/5000-15/G062 (Grad +2)

Resultaten är alltså minst 100 gånger sannolikare om skoavtrycket S060 är avsatt med vänsterskon S062 än om det är avsatt med en annan sko.

Polisen

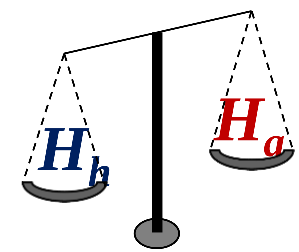
H_h : Beslagtagen sko G har avsatt det säkrade skoavtrycket S

H_a : En annan sko har avsatt spåret

Annan information i utredningen

Hur sannolik utifrån annan information i utredningen (på förhand) är…

H_h : Beslagtagen sko G har avsatt det säkrade skoavtrycket S $\Rightarrow P(H_h)$



…och hur sannolik utifrån annan information i utredningen (på förhand) är…

H_a : En annan sko har avsatt spåret $\Rightarrow P(H_a)$

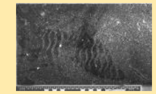
$$\text{Föroddsen} = O = \frac{P(H_h)}{P(H_a)}$$

Anger hur mycket som på förhand talar för huvudhypotesen i förhållande till hur mycket som på förhand talar för alternativhypotesen.

H_h **:** Beslagtagen sko G har avsatt det säkrade skoavtrycket S

H_a **:** En annan sko har avsatt spåret

Ett säkrat skoavtryck och ett par beslagtagna skor.



Annan information i utredningen

$$O = \frac{P(H_h)}{P(H_a)}$$

$$V = \frac{P(\text{Resultaten} \mid H_h)}{P(\text{Resultaten} \mid H_a)}$$

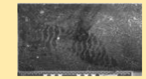
Forensisk bedömning (*vid laboratoriet*)

$$\text{Bayes' sats}: \quad \frac{P(H_h \mid \text{Resultaten})}{P(H_a \mid \text{Resultaten})} = V \times O$$

Efteroddsen

$$P(H_h \mid \text{Resultaten})$$

Slutlig sannolikhet för H_h
- Bevisvärdet (tilltron)

Ett säkrat skoavtryck och ett par beslagtagna skor.  

Säg att **förundersökningsledaren** har bedömt föroddsen för att en av här skorna skulle ha avsatt avtrycket till 2 mot 1, dvs.

$$O = \frac{P(H_h)}{P(H_a)} \approx 2$$

Innebär att $P(H_h)$ initialt bedöms vara c:a 0,67

Innan resultat från den forensiska undersökningen har beaktats är det ungefär dubbelt så sannolikt att en av skorna har avsatt avtrycket jämfört med att någon annan sko har gjort det.

Från sakkunnigutlåtandet från NFC:

$$\Rightarrow V = \frac{P(\text{Resultaten} \mid H_h)}{P(\text{Resultaten} \mid H_a)} > 100$$

Resultaten är minst 100 gånger sannolikare om skoavtrycket S060 är avsatt med vänsterskon G062 än om det är avsatt med en annan sko.

$$\Rightarrow \frac{P(H_h \mid \text{Resultaten})}{P(H_a \mid \text{Resultaten})} = V \times O > 100 \times 2 = 200$$

Detta ger att den slutliga sannolikheten för att H_h stämmer, dvs. att det var vänsterskon G062 som avsatte avtrycket S060 blir

$$P(H_h \mid \text{Resultaten}) = \frac{200}{200+1} \approx 0{,}995$$

Polisen

H_h **:** Beslagtagen sko G har avsatt det säkrade skoavtrycket S

H_a **:** En annan sko har avsatt spåret

Bevisvärde för vem?...

En källnivåjämförelse är som regel en forensisk undersökning med förundersökningsledaren (oftast en åklagare) som ”destination”.

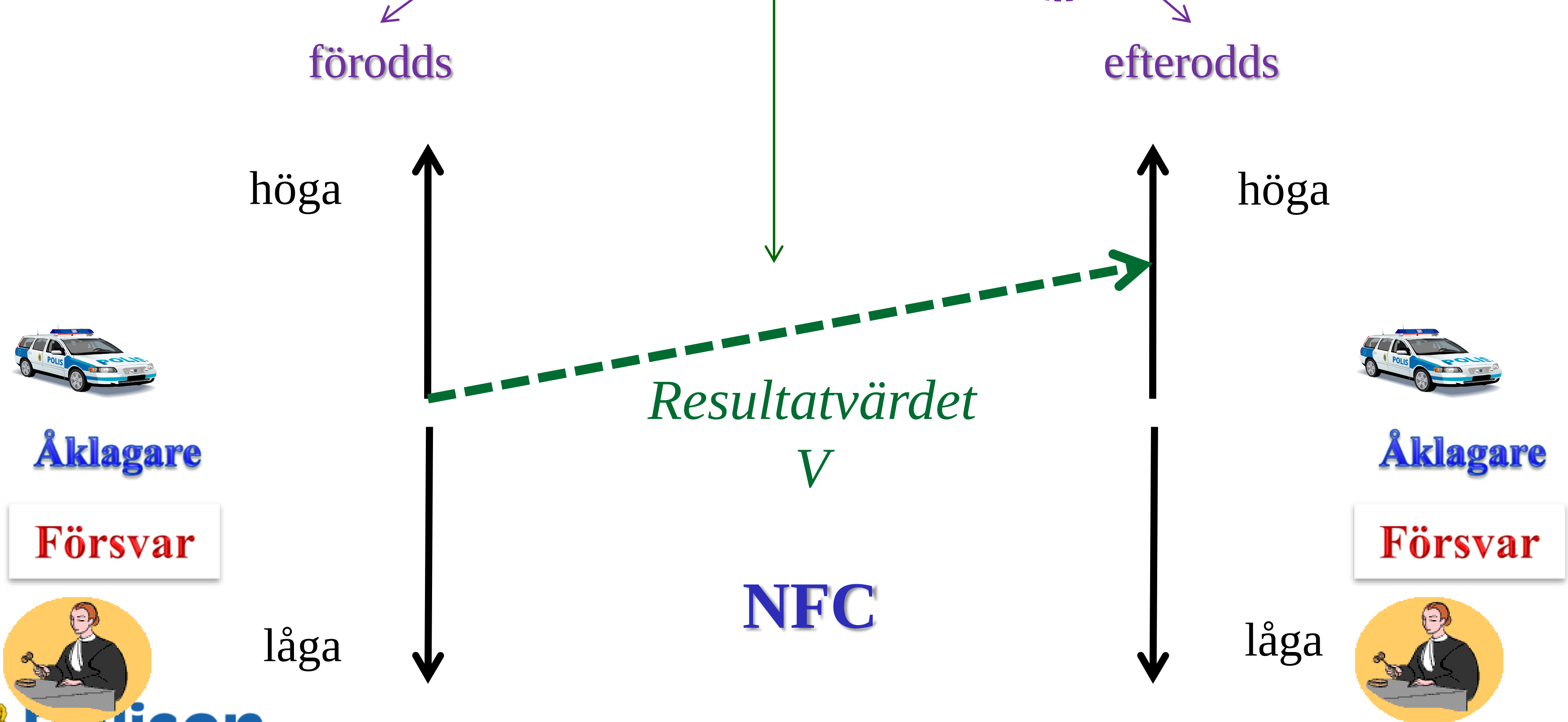
Det är förundersökningsledaren som äger frågeställningen och som (i teorin)…

- är inblandad i valet av alternativhypotes
- har att bedöma storleken hos föroddsen
- ska ta ställning till:

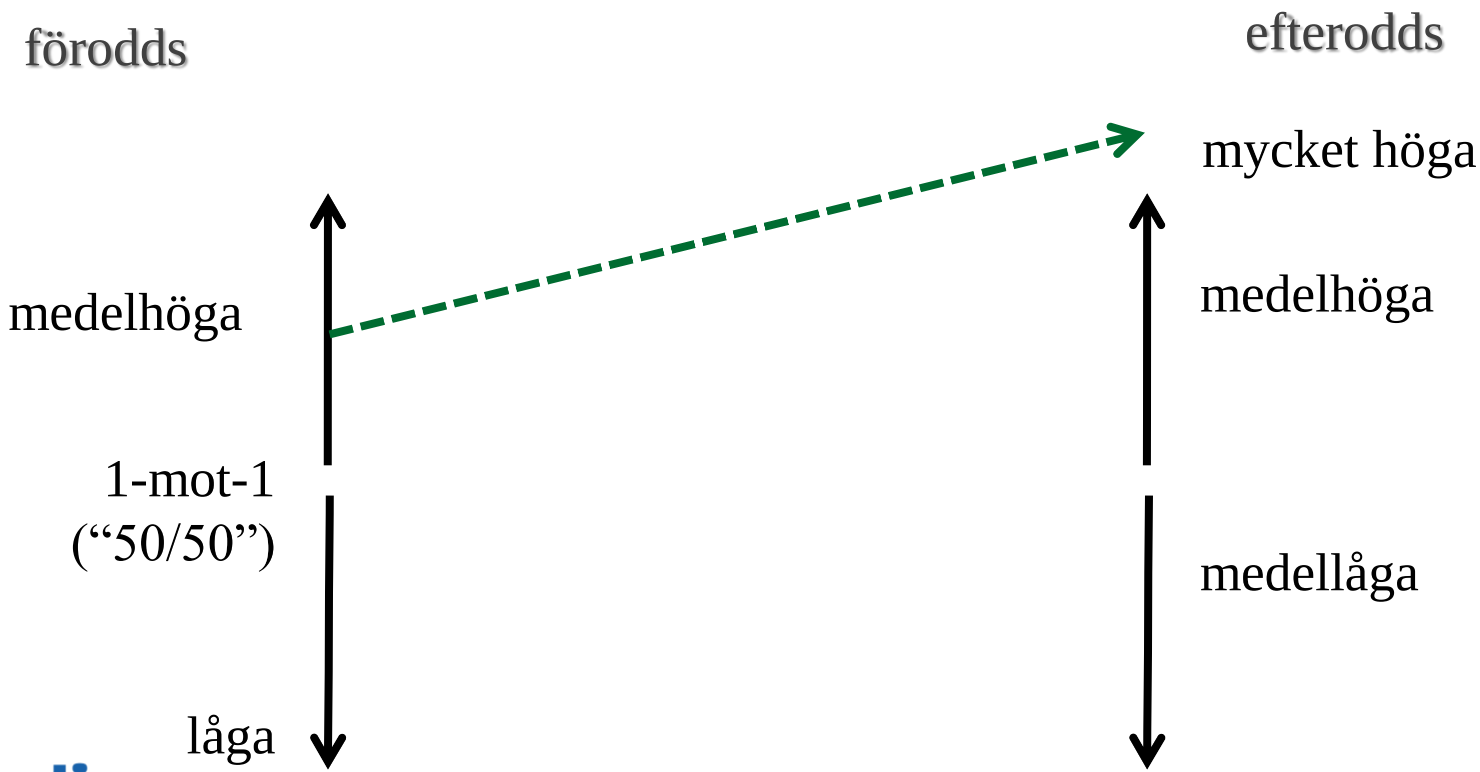
Räcker bevisvärdet för att ta med materialet som bevis i åtalet?

Grafisk illustration av Bayes' sats

$$\frac{P(H_h)}{P(H_a)} \times \frac{P(E|H_h)}{P(E|H_a)} = \frac{P(H_h|E)}{P(H_a|E)}$$



- Med samma par av huvud- och alternativhypotes har ett resultat <u>alltid samma resultatvärde</u> dvs. en och samma vinkel
- Däremot kan vi få olika <u>bevisvärden</u> (efterodds) beroende på hur föroddsen ser ut



Polisen

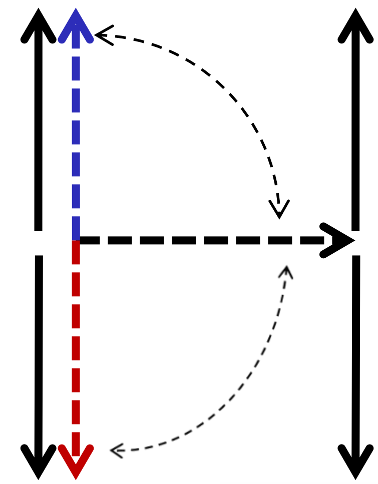
Varför ser intervallen i skalan ut som de gör?

Slutsats-grad	Resultatvärdets storlek	”Förklaring”
		Resultaten är…
+4	minst en miljon	…minst en miljon ggr sannolikare…
+3	mellan 6000 och en miljon	…minst 6000 ggr sannolikare…
+2	mellan 100 och 6000	…minst 100 ggr sannolikare…
+1	mellan 6 och 100	…minst 6 ggr sannolikare…
0	mellan 1/6 och 6	… ungefär lika sannolika…
		…om huvudhypotesen är sann jämfört med om den alternativa hypotesen är sann
–1	mellan 1/100 och 1/6	…minst 6 ggr sannolikare…
–2	mellan 1/6000 och 1/100	…minst 100 ggr sannolikare…
–3	mellan 1/(en miljon) och 1/6000	…minst 6000 ggr sannolikare…
–4	högst 1/(en miljon)	…minst en miljon ggr sannolikare…
		…om den alternativa hypotesen är sann jämfört med om den huvudhypotesen är sann

Polisen

- Resultatvärdet är en kvot mellan två sannolikheter

$$\text{Resultatvärdet} = V = \frac{P(\text{Resultaten}|H_h)}{P(\text{Resultaten}|H_a)} = \frac{P(E|H_h)}{P(E|H_a)}$$



- I teorin kan resultatvärdet variera från att
  - vara lika med noll (***huvudhypotesen*** utesluts)

  till att
  - vara oändligt stort (***alternativhypotesen*** utesluts)

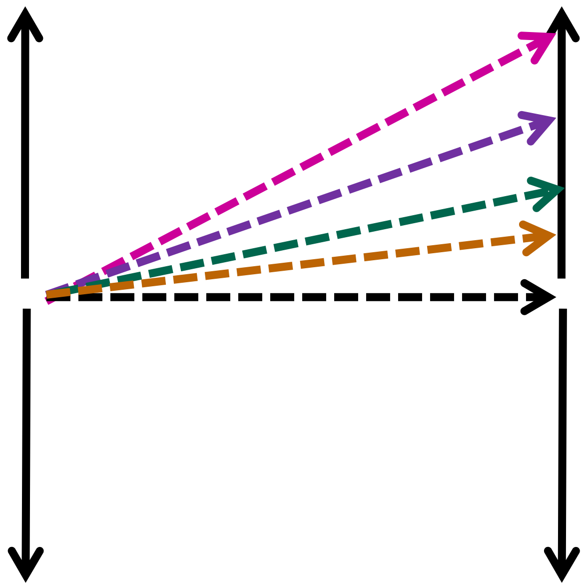
- …men en användbar skala måste ha ett mycket begränsat antal skalsteg – dela in ett oändligt långt intervall i ett ändligt antal delintervall

- NFC har valt en skala *symmetrisk* kring resultatvärdet = 1 och med fyra positiva (stödjande) och fyra negativa (icke stödjande) skalsteg, vart och ett motsvarande ett intervall av resultatvärden

- Intervallgränserna har valts så att *bevisvärdet* blir acceptabelt högt utifrån skalstegets nivå *när föroddsen är 1 mot 1* , dvs. när de två hypoteserna på förhand bedöms vara lika sannolika

Slutsats-grad
+4
+3
+2
+1
0
–1
–2
–3
–4

Förodds:
"1 mot 1"



Grad	Bevisvärde $P(H_h \mid E)$	Resulterande nedre gräns för V
+4 :	**> 0,999999**	**1 miljon ≤ *V***
+3:	**> 0,9998**	**6000 ≤ *V***
+2:	**> 0,99**	**100 ≤ *V***
+1:	**> 0,86**	**6 ≤ *V***
0:	**mellan 0,14 och 0,86**	**(1/6 < *V* < 6)**

Grad + 2: Sannolikheten 0,99 (= 99 %) är i rättssammanhang allmänt vedertagen som tillräckligt hög för att "styrka" en hypotes (att ett enskilt resultat tas med i bevisningen)

Grad + 4: Gränsen 1 miljon är ett historiskt "arv" från hur +4 har valts för resultatvärden för dna-överensstämmelser

Grad +1 och + 3: Intervallen av resultatvärden har valts så att de successivt ökar i längd på ett matematiskt regelbundet sätt. Sannolikheterna 0,9998 och 0,86 faller då automatiskt ut

Nordgaard A., Ansell R., Drotz W. & Jaeger L.: "Scale of conclusions for the value of evidence". *Law, Probability and Risk* 11(1): 1-24.

Resultatvärdering och IT-forensik?

- Till stor del ”under utveckling”, men…

- Bildanalys vid NFC: I framkant när det gäller resultatvärdering vid NFC
  - Tidig med implementering av standardmässiga metoder för bildjämförelser
  - Driver utveckling inom ansiktsjämförelseområdet – validering av algoritmbaserade metoder
  - …

Datorer och mobiltelefoner?

- Inga självklara frågeställningar från polisutredningen
- Är mer upp till den forensiska expertisen att formulera frågorna – *Investigative mode* (utredande fas)
- Likheter med brottsplatsanalys och (!!) rättspatologi (obduktioner)

Mer ingående om jämförelser

Hypotesformulering vid jämförelser av material

H_h : Omstritt material och jämförelsematerial har samma ursprung
H_a : Omstritt material och jämförelsematerial har olika ursprung

”Glasbiten kommer från rutan” vs. ”Glasbiten kommer från ett annat glasföremål”

”Amfetaminbeslagen kommer från samma utfällningssats” vs. ”Amfetaminbeslagen kommer från olika utfällningssatser”

”Avtrycket har avsatts med denna sko” vs. ”Avtrycket har avsatts med en annan sko”

”Överkroppsplagget på bilden är den misstänktes jacka” vs. ”Överkroppsplagget på bilden är ett annat överkroppsplagg”

Resultat, ***E***

- Materialbedömning av omstritt material och jämförelsematerial – Skapar delvis förutsättningar för värderingen
- Observationer på omstritt material och jämförelsematerial
- Resultat av kemisk/biologisk/teknisk analys av omstritt material och jämförelsematerial
- Utfall av jämförelse mellan omstritt material och jämförelsematerial med avseende på gjorda observationer och resultat från kemisk/biologisk/teknisk analys

Om nu t.ex. jämförelsen har gett en ”överensstämmelse”, är då

$$\frac{P(E \mid H_h)}{P(E \mid H_a)} \quad \text{detsamma som} \quad \frac{P(\text{"Överensstämmelse"} \mid H_h)}{P(\text{"Överensstämmelse"} \mid H_a)} \quad ?$$

Dela upp resultatet från observationer och analys i

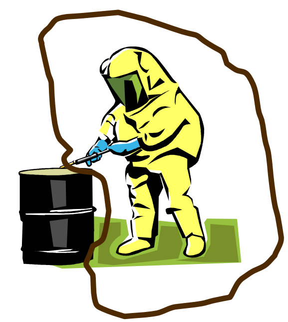
E_o = Observationer och analysresultat för omstritt material
E_j = Observationer och analysresultat för jämförelsematerial

Låt $E_{obs} = (E_o, E_j)$

Låt vidare

M_o = Omständigheter runt omstritt material (inkl. materialbedömning)
M_j = Omständigheter runt jämförelsematerial (inkl. materialbedömning)

Omständigheter runt materialen har inverkan på vad det är man observerar och påverkar därför sannolikheten att man ser just det man ser. *Ex. dålig bildkvalitet, smutsiga skoavtryck*



Låt $E_M = (M_o, M_j)$

$$E = (E_{obs}, E_M)$$

E_{o} = Observationer och analysresultat för omstritt material
E_{j} = Observationer och analysresultat för jämförelsematerial
M_{o} = Omständigheter runt omstritt material
M_{j} = Omständigheter runt jämförelsematerial

Resultatvärdet där såväl observationer och analysresultat som material-omständigheter inräknats kan då skrivas

$$P(A,B) = P(A|B) \times P(B)$$
$$P(A,B|C) = P(A|B,C) \times P(B|C)$$

$$V = \frac{P(E|H_h)}{P(E|H_a)} = \frac{P(E_{obs}, E_M|H_h)}{P(E_{obs}, E_M|H_a)} = \frac{P(E_{\text{o}}, E_{\text{j}}, M_{\text{o}}, M_{\text{j}}|H_h)}{P(E_{\text{o}}, E_{\text{j}}, M_{\text{o}}, M_{\text{j}}|H_a)} = \frac{P(E_{\text{o}}, E_{\text{j}}|M_{\text{o}}, M_{\text{j}}, H_h) \times P(M_{\text{o}}, M_{\text{j}}|H_h)}{P(E_{\text{o}}, E_{\text{j}}|M_{\text{o}}, M_{\text{j}}, H_a) \times P(M_{\text{o}}, M_{\text{j}}|H_a)} =$$

$$= \frac{P(E_{\text{o}}, E_{\text{j}}|M_{\text{o}}, M_{\text{j}}, H_h)}{P(E_{\text{o}}, E_{\text{j}}|M_{\text{o}}, M_{\text{j}}, H_a)} \times \boxed{\frac{P(M_{\text{o}}, M_{\text{j}}|H_h)}{P(M_{\text{o}}, M_{\text{j}}|H_a)}} = V_{obs} \times V_M$$

Den andra kvoten i V, dvs. V_M, har NFC svårt att bedöma (om det ens kan/ska göras), dvs hur sannolika materialomständigheterna är under respektive hypotes.

Ofta går det dock att sätta kvoten till ett, dvs. vi kan anta att sannolikheterna i täljare och nämnare är lika stora, *dvs.* vi antar att hypoteserna inte påverkar materialomständigheterna.

Detta ger förenklingen

$$V = \frac{P(E_{\text{o}}, E_{\text{j}}|M_{\text{o}}, M_{\text{j}}, H_h)}{P(E_{\text{o}}, E_{\text{j}}|M_{\text{o}}, M_{\text{j}}, H_a)}$$

$$V = \frac{P\left(E_o, E_j \middle| M_o, M_j, H_h\right)}{P\left(E_o, E_j \middle| M_o, M_j, H_a\right)}$$

E_o = Observationer och analysresultat för omstritt material
E_j = Observationer och analysresultat för jämförelsematerial
M_o = Omständigheter runt omstritt material
M_j = Omständigheter runt jämförelsematerial

Exempel

Vi har en bild från en övervakningskamera i vilket ansiktet på en person syns. Detta ansikte ska jämföras med fotografier av en person som misstänks vara personen på bilden (H_h).


Låt t.ex.
E_o = ”En synlig fläck på högra kinden på ansiktet på bilden”
E_j = ”En observerad leverfläck på personens på fotografiet vänstra kind”

M_o = ”Bilden från övervakningskameran har låg kvalitet”
M_j = ”Fotografierna är tagna senare än bilden från övervakningskameran”

$P(E_o, E_j \mid M_o, M_j, H_h)$ är inte med nödvändighet hög här. Dålig bildkvalitet gör att vi inte säkert kan se leverfläcken på övervakningsbilden (samtidigt som vi ser den på fotografierna) även om vi utgår från att personerna på bild och fotografi är en och samma.

$P(E_o, E_j \mid M_o, M_j, H_a)$ kan vara relativt hög. I en bild med dålig kvalitet kan mycket observeras som kanske inte ens finns på den person som avbildats.

$$V=\frac{P(E_o,E_j|M_o,M_j,H_h)}{P(E_o,E_j|M_o,M_j,H_a)}$$

E_o = Observationer och analysresultat för omstritt material
E_j = Observationer och analysresultat för jämförelsematerial
M_o = Omständigheter runt omstritt material
M_j = Omständigheter runt jämförelsematerial

Materialomständigheterna (M_o och M_j) är viktiga komponenter, men för att göra framställningen enklare bortser vi från att ta med dem i formlerna.

$$V=\frac{P(E_o,E_j|\ldots,H_h)}{P(E_o,E_j|\ldots,H_a)}$$

Utveckla vidare:

$$V=\frac{P(E_o,E_j|H_h)}{P(E_o,E_j|H_a)}=\frac{P(E_o|E_j,H_h)\times P(E_j|H_h)}{P(E_o|E_j,H_a)\times P(E_j|H_a)}$$

Oftast beror inte E_j av vilken hypotes som är sann – t.ex. har en misstänkt sitt dna oavsett om hen har lämnat ett spår eller ej; en glasruta har sin grundämnessammansättning oavsett om ett omstritt fragment kommer från den eller ej.

$$\Rightarrow \frac{P(E_j|H_h)}{P(E_j|H_a)} \text{ kan sättas till } 1 \quad \Rightarrow \quad V=\frac{P(E_o|E_j,H_h)}{P(E_o|E_j,H_a)}$$

Polisen

$$V = \frac{P(E_o | E_j, H_h)}{P(E_o | E_j, H_a)}$$

$\boldsymbol{H_h}$: Omstritt material och jämförelsematerial har samma ursprung
$\boldsymbol{H_a}$: Omstritt material och jämförelsematerial har olika ursprung
E_o = Observationer och analysresultat för omstritt material
E_j = Observationer och analysresultat för jämförelsematerial

Vidare, om H_a är sann så beror inte E_o av E_j – egenskaperna hos omstritt material har i detta läge inte med jämförelsematerialet att göra.

$$\Rightarrow V = \frac{P(E_o | E_j, H_h)}{P(E_o | H_a)}$$

Täljaren kommer att omfatta sannolikheter för överensstämmelser och sannolikheter för avvikelser utgående från att material har samma ursprung.
Inget som förväntas finnas i några databaser.

Nämnaren anger hur vanliga observationerna och analysresultaten för omstritt material (E_o) är i allmänhet. Information, som typiskt kan fås fram ur ”databaser”.

Exemplet med bildjämförelsen

H_h : Det är den misstänktes ansikte som syns på övervakningsbilden
H_a : Det är någon annans persons ansikte som syns…
E_o = ”En synlig fläck på vänstra kinden på ansiktet på bilden”
E_j = ”En observerad leverfläck på den misstänktes vänstra kind”

Formeln $V = \dfrac{P\left(E_o \middle| E_j, H_h\right)}{P\left(E_o \middle| H_a\right)}$ skapar naturliga förutsättningar för värderingen.

$P(E_o \mid E_j, H_h)$ anger hur sannolikt/förväntat det är att se en fläck på vänster kind på personen på bilden <u>om</u> det är personen på fotografiet <u>och</u> fotografiet visar på en leverfläck på vänster kind.
Denna sannolikhet påverkas förstås av materialomständigheterna (även om vi inte har tagit med dessa i formeln just här)

$P(E_o \mid H_a)$ anger hur ofta vi skulle observera en fläck på vänster kind på en person på en övervakningsbild. Denna sannolikhet har inget med personen på fotografiet att göra, men väl med materialomständigheterna. Här måste vi väga in både förekomsten av fläckar på personer och hur sannolikt det är att de observeras [vid rådande kvalitet].

Polisen

$$V=\frac{P\left(E_o \middle| E_j, H_h\right)}{P\left(E_o \middle| H_a\right)}$$

$\boldsymbol{H_h}$: Omstritt material och jämförelsematerial har samma ursprung
$\boldsymbol{H_a}$: Omstritt material och jämförelsematerial har olika ursprung
E_o = Observationer och analysresultat för omstritt material
E_j = Observationer och analysresultat för jämförelsematerial

Om E_o och E_j är klasskaraktäristika, dvs. anges som distinkta värden eller tillstånd – *t.ex.* upplevd färgnyans, genotyper, tryckteknik, fiberfabrikat, sulmönster, tablettmärkning etc.

så innebär en *överensstämmelse* – teoretiskt – att $E_o = E_j$ och en *avvikelse* att $E_o \neq E_j$

Om H_h är sann förväntar vi oss därför att $E_o = E_j$, dvs. $E_j = e \rightarrow E_o = e$

Materialomständigheterna kan dock påverka så att

$P(E_o = e \mid E_j = e, H_h)$ inte självklart = 1 (och därmed att $P(E_o \neq e \mid E_j = e, H_h)$ inte självklart = 0)

Om E_o och E_j är karaktäristika, som varierar kontinuerligt – *t.ex.* brytningsindex, konduktivitet, toppareor i kromatogram, längd, vikt

så kan <u>inte</u> ens i teorin $P(E_o = e \mid E_j = e, H_h)$ vara = 1

Polisen

I mån av tid…

Den kontinuerliga varianten vid jämförelser

När det vi observerar varierar kontinuerligt är det mer praktiskt att använda traditionella variabelsymboler (x, y, …).

När vi mäter något som varierar kontinuerligt är det också normalt att vi gör mer än en mätning.

Låt därför E_j representeras av $\mathbf{x} = (x_1, x_2, \ldots, x_m)$ som betecknar de m mätvärden som fås från *jämförelsematerialet*

och

E_o representeras av $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ som betecknar de n mätvärden som fås från *omstritt material*.

Obs! Vi utgår från att vi faktiskt kan göra olika många mätningar på respektive material – t.ex. på grund av att alla planerade mätningar inte ger lyckat resultat.

Polisen

Typiska värden på m och n vid NFC är 2, 3, 4 och 5

$$V = \frac{P\left(E_\text{o} \middle| E_\text{j}, H_h\right)}{P\left(E_\text{o} \middle| H_a\right)} \quad \text{motsvaras nu av} \quad V = \frac{f\left(\mathbf{y} \middle| \mathbf{x}, H_h\right)}{f\left(\mathbf{y} \middle| H_a\right)}$$

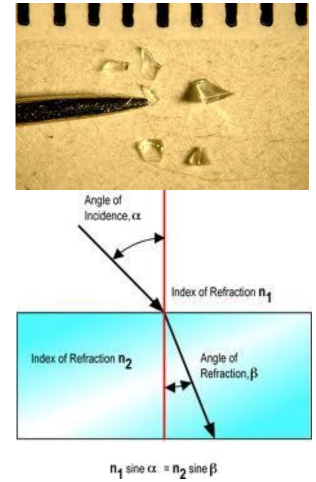
där $f\,(\mathbf{y} \mid \mathbf{x}\,,\, H_h\,)$ är den betingade täthetsfunktionens värde i de observerade värdena **y** på omstritt material, dvs. $(y_1, \ldots, y_n)$, under antagande att H_h är sann och givet de observerade värdena **x** på jämförelsematerialet, dvs. $(x_1, \ldots, x_m)$,

och $f\,(\mathbf{y} \mid H_a\,)$ är (den betingade) täthetsfunktionens värde i de observerade värdena **y** på omstritt material under antagande att H_a är sann.

$f\,(\mathbf{y} \mid H_a\,)$ motsvarar dock den täthetsfunktion som allmänt gäller för den typ av observationer vi gör på det omstridda materialet. – jfr. databaser från vilka vi beräknar relativa frekvenser.

Polisen

Exempel: Brytningsindex hos glas



Ett omstritt glasfragment jämförs med en krossad ruta

H_h : Glasfragmentet kommer från den krossade rutan
H_a : Glasfragmentet kommer från något annat glasföremål

E : ”Överensstämmelse” i brytningsindex, *RI* (värdena ligger mycket nära)

(Vi begränsar oss i detta exempel till uppmätt brytningsindex före s.k. annealing)

I mer detalj:

Vi har gjort tre mätningar av brytningsindex på vardera glas:

Omstritt fragment: $\mathbf{y} = (1{,}51592,\ 1{,}51591,\ 1{,}51594)$

Jämförelseglas från ruta: $\mathbf{x} = (1{,}51591,\ 1{,}51591,\ 1{,}51593)$

Hur får vi nu fram täljare och nämnare i

$$V = \frac{f(\mathbf{y}|\mathbf{x}, H_h)}{f(\mathbf{y}|H_a)} \qquad ?$$

H_h : Glasfragmentet kommer från den krossade rutan
H_a : Glasfragmentet kommer från något annat glasföremål
E: y = (1,51592, 1,51591, 1,51594)
x = (1,51591, 1,51591, 1,51593)

1. Normalfördelad variation överallt kan antas

Två viktiga komponenter:

- Spridning mellan mätningar på ett och samma objekt eller mellan mätningar på olika objekt med ***samma*** ursprung – *inomvariation*

  Representeras av en varians σ^2

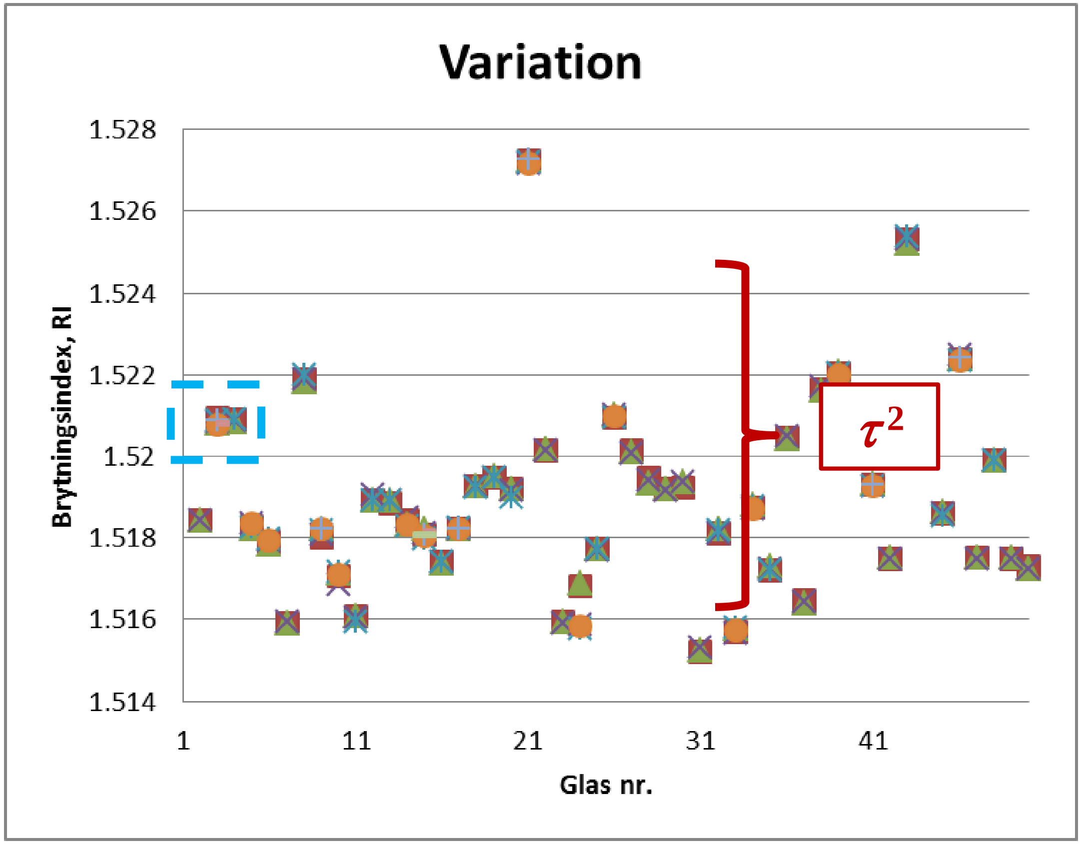
- Spridning i faktisk egenskap (här: brytningsindex) mellan objekt med ***olika*** ursprung – *mellanvariation*

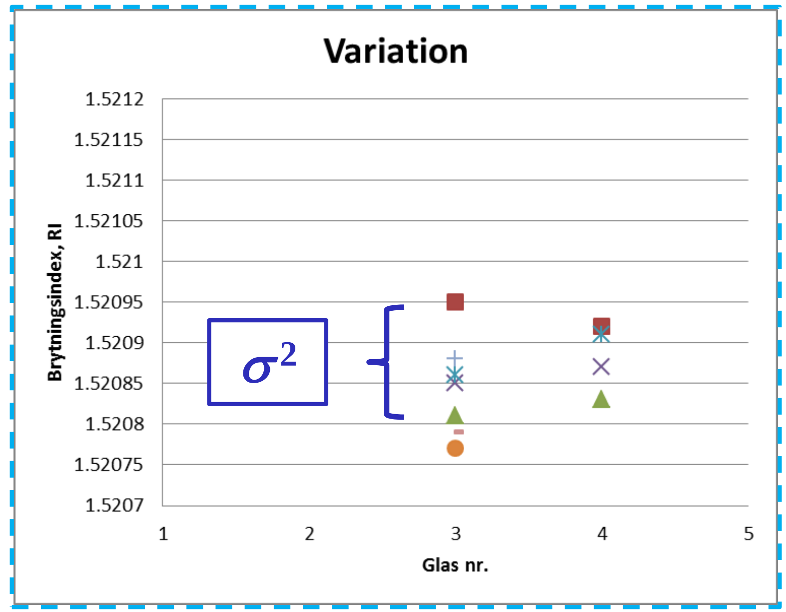
  Representeras av en annan varians τ^2

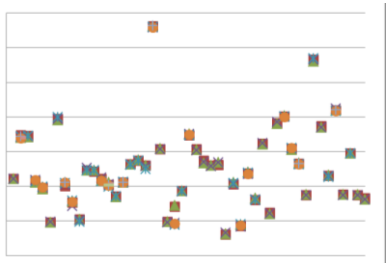
I allmänhet är τ^2 mycket större än σ^2 $\qquad (\tau^2 >> \sigma^2)$

H_h : Glasfragmentet kommer från den krossade rutan
H_a : Glasfragmentet kommer från något annat glasföremål
E: $\mathbf{y} = (1{,}51592,\ 1{,}51591,\ 1{,}51594)$
$\mathbf{x} = (1{,}51591,\ 1{,}51591,\ 1{,}51593)$

Uppmätta brytningsindex hos 50 olika glas, minst tre mätningar på varje glas







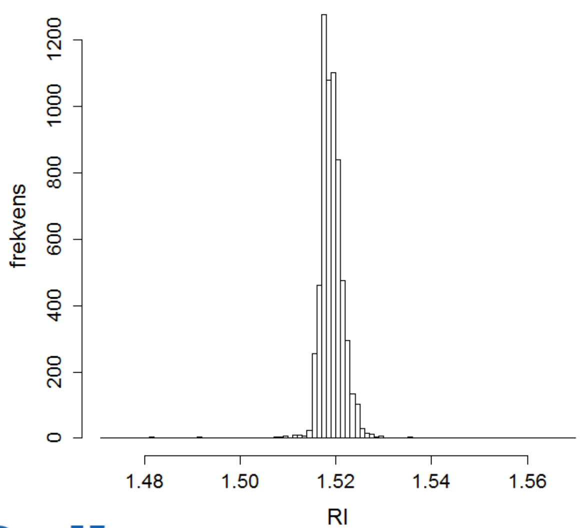
$\Big\}\ \tau^2$ $\}\ \sigma^2$

H_h : Glasfragmentet kommer från den krossade rutan
H_a : Glasfragmentet kommer från något annat glasföremål
E: $\mathbf{y}$ = (1,51592, 1,51591, 1,51594)
$\mathbf{x}$ = (1,51591, 1,51591, 1,51593)

Från upprepade mätningar på 50 glas kan vi få en stabil uppskattning av σ^2 – här $1{,}08 \cdot 10^{-8}$.

Hur uppskattar vi storleken hos τ^2 ?

NFC:s glasdatabas: 6185 glas per 2018-05-08



Kan vid en första anblick se ut som normalfördelad variation, men detta stämmer inte så bra. Formen är för toppig och kanske med väl korta svansar.

Ger hur som helst en bra skattning av τ^2 : $1{,}003 \cdot 10^{-5}$

Polisen

Hur utnyttjar vi nu antagandet om normalfördelad variation (överallt)?

H_h : Glasfragmentet kommer från den krossade rutan
H_a : Glasfragmentet kommer från något annat glasföremål
E: **y** = (1,51592, 1,51591, 1,51594)
x = (1,51591, 1,51591, 1,51593)

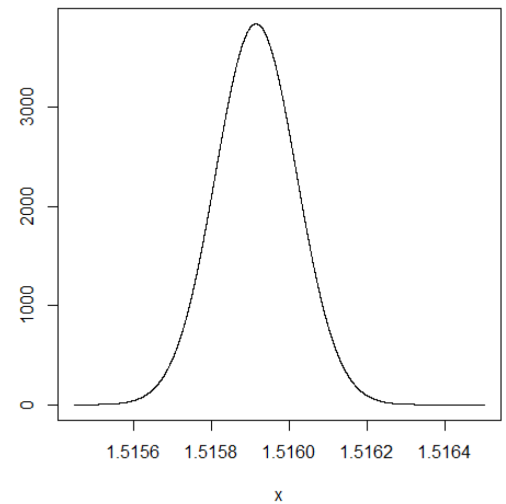
Jämförelseglaset från rutan har ett ”sant” brytningsindex, kalla detta RI_{j}

Mätvärden på jämförelseglaset, kallad variabeln X, har då en normalfördelad variation med genomsnittsvärde (väntevärde) RI_{j} och standardavvikelse σ (varians σ^2) .

Internationellt skrivsätt: $(X \mid RI_{\mathrm{j}} , \sigma^2) \sim N(RI_{\mathrm{j}} , \sigma^2)$

Täthetsfunktion:

$$f_X\left(x \middle| RI_{\mathrm{j}}, \sigma^2\right) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{\frac{\left(x - RI_{\mathrm{j}}\right)^2}{2 \cdot \sigma^2}}$$



H_h : Glasfragmentet kommer från den krossade rutan
H_a : Glasfragmentet kommer från något annat glasföremål
E: $\mathbf{y} = (1{,}51592,\ 1{,}51591,\ 1{,}51594)$
$\mathbf{x} = (1{,}51591,\ 1{,}51591,\ 1{,}51593)$

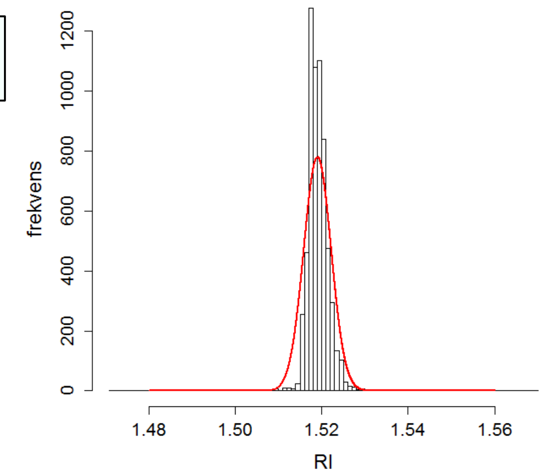
Motsvarande har fragmentet ett ”sant” brytningsindex, kalla detta RI_{o}

Mätvärden på fragmentet, kallad variabeln Y, har då en normalfördelad variation med genomsnittsvärde RI_{o} och standardavvikelse σ – dvs. lika spridning som X.

$(Y \mid RI_{\mathrm{o}}, \sigma^2) \sim N(RI_{\mathrm{o}}, \sigma^2)$

Men vi vet ju inte vad vare sig RI_{j} eller RI_{o} är – vi skall i första hand undersöka om de är lika eller ej – *En slags sammansatt huvudhypotes*

<u>Vikta</u> över variationen i ”sanna” brytningsindex via glasdatabasen



Med fortsatt antagande om normalfördelning (överallt):

$$\left(RI_{\mathrm{j}} \,\middle|\, \mu, \tau^2\right) \sim N\left(\mu, \tau^2\right)$$
$$\left(RI_{\mathrm{o}} \,\middle|\, \mu, \tau^2\right) \sim N\left(\mu, \tau^2\right)$$

Ger uppskattade värden $\mu \approx 1{,}51909$ och (som tidigare) $\tau^2 \approx 1{,}003 \cdot 10^{-5}$

H_h : Glasfragmentet kommer från den krossade rutan
H_a : Glasfragmentet kommer från något annat glasföremål
E: $\mathbf{y} = (1{,}51592,\ 1{,}51591,\ 1{,}51594)$
$\mathbf{x} = (1{,}51591,\ 1{,}51591,\ 1{,}51593)$

Vi har alltså täthetsfunktioner för X och Y :

$$f_X\left(x \mid RI_{\mathrm{j}}, \sigma^2\right) \qquad f_Y\left(y \mid RI_{\mathrm{o}}, \sigma^2\right)$$

men också för variationen i ”sant” brytningsindex: $f_{RI}\left(r \mid \mu, \tau^2\right) = \dfrac{1}{\tau \cdot \sqrt{2\pi}} \cdot e^{-\frac{(r-\mu)^2}{2 \cdot \tau^2}}$

över hela populationen av glasföremål.

En viktning skapar nu s.k. *marginella* täthetsfunktioner för X och Y:

$$f_X\left(x \mid \mu, \sigma^2, \tau^2\right) = \int f_X\left(x \mid r, \sigma^2\right) \cdot f_{RI}\left(r \mid \mu, \tau^2\right) dr$$
$$f_Y\left(y \mid \mu, \sigma^2, \tau^2\right) = \int f_Y\left(y \mid r, \sigma^2\right) \cdot f_{RI}\left(r \mid \mu, \tau^2\right) dr$$

Viktning med en kontinuerlig viktfunktion görs med en integral.

som visar sig ge att

$(X \mid \mu, \sigma^2, \tau^2) \sim N(\mu, \sigma^2 + \tau^2)$ och
$(Y \mid \mu, \sigma^2, \tau^2) \sim N(\mu, \sigma^2 + \tau^2)$

Med marginalfördelning menar vi variationen generellt i ett uppmätt brytningsindex, *utan* att betinga på dess ursprung

I praktiken intresserar vi oss förstås inte för ett enda observerat värde utan vi tar medelvärden $(\bar{y}, \bar{x})$ av gjorda mätningar.

H_h : Glasfragmentet kommer från den krossade rutan
H_a : Glasfragmentet kommer från något annat glasföremål
E: $\mathbf{y} = (1{,}51592,\ 1{,}51591,\ 1{,}51594)$
$\mathbf{x} = (1{,}51591,\ 1{,}51591,\ 1{,}51593)$

$$\bar{y} = (1{,}51592 + 1{,}51591 + 1{,}51594)/3 \approx 1{,}515923$$
$$\bar{x} = (1{,}51591 + 1{,}51591 + 1{,}51593)/3 \approx 1{,}515917$$

Vi har utgått från *m* mätvärden på jämförelsematerialet och *n* mätvärden på omstritt material – i detta exempel är $m = n = 3$.

Variationen blir dock fortfarande normalfördelad, men varianserna förändras:

Specifikt för ett visst glas: Generellt sett över alla sorters glas:

$$\left(\bar{X} \mid RI_{\mathrm{j}}, \sigma^2\right) \sim N\left(RI_{\mathrm{j}}, \frac{\sigma^2}{m}\right) \quad ; \quad \left(\bar{X} \mid \mu, \sigma^2, \tau^2\right) \sim N\left(\mu, \frac{\sigma^2}{m} + \tau^2\right)$$

$$\left(\bar{Y} \mid RI_{\mathrm{o}}, \sigma^2\right) \sim N\left(RI_{\mathrm{o}}, \frac{\sigma^2}{n}\right) \quad ; \quad \left(\bar{Y} \mid \mu, \sigma^2, \tau^2\right) \sim N\left(\mu, \frac{\sigma^2}{n} + \tau^2\right)$$

Statistikteoretiskt: Med kända varianser (σ^2 och τ^2) innehåller medelvärdena $(\bar{x}, \bar{y})$ ***tillräcklig*** information om sanna brytningsindex.

$$\left(\bar{X} \mid RI_j, \sigma^2\right) \sim N\left(RI_j, \frac{\sigma^2}{m}\right) \quad ; \quad \left(\bar{X} \mid \mu, \sigma^2, \tau^2\right) \sim N\left(\mu, \frac{\sigma^2}{m} + \tau^2\right)$$

$$\left(\bar{Y} \mid RI_o, \sigma^2\right) \sim N\left(RI_o, \frac{\sigma^2}{n}\right) \quad ; \quad \left(\bar{Y} \mid \mu, \sigma^2, \tau^2\right) \sim N\left(\mu, \frac{\sigma^2}{n} + \tau^2\right)$$

H_h : Glasfragmentet kommer från den krossade rutan
H_a : Glasfragmentet kommer från något annat glasföremål
E: y = (1,51592, 1,51591, 1,51594)
x = (1,51591, 1,51591, 1,51593)

A. Approximativ variant (Evett, 1986)

Under antagande att H_h är sann – vilket betyder att **y** förväntas ligga nära **x** – och givet värdena **x** antar vi modellen

$$(\bar{Y} | \bar{x}, \sigma^2, H_h) \sim N\left(\bar{x}, \frac{\sigma^2}{n}\right) \quad ; \quad (\bar{Y} | \mu, \sigma^2, \tau^2, H_a) \sim N\left(\mu, \frac{\sigma^2}{n} + \tau^2\right)$$

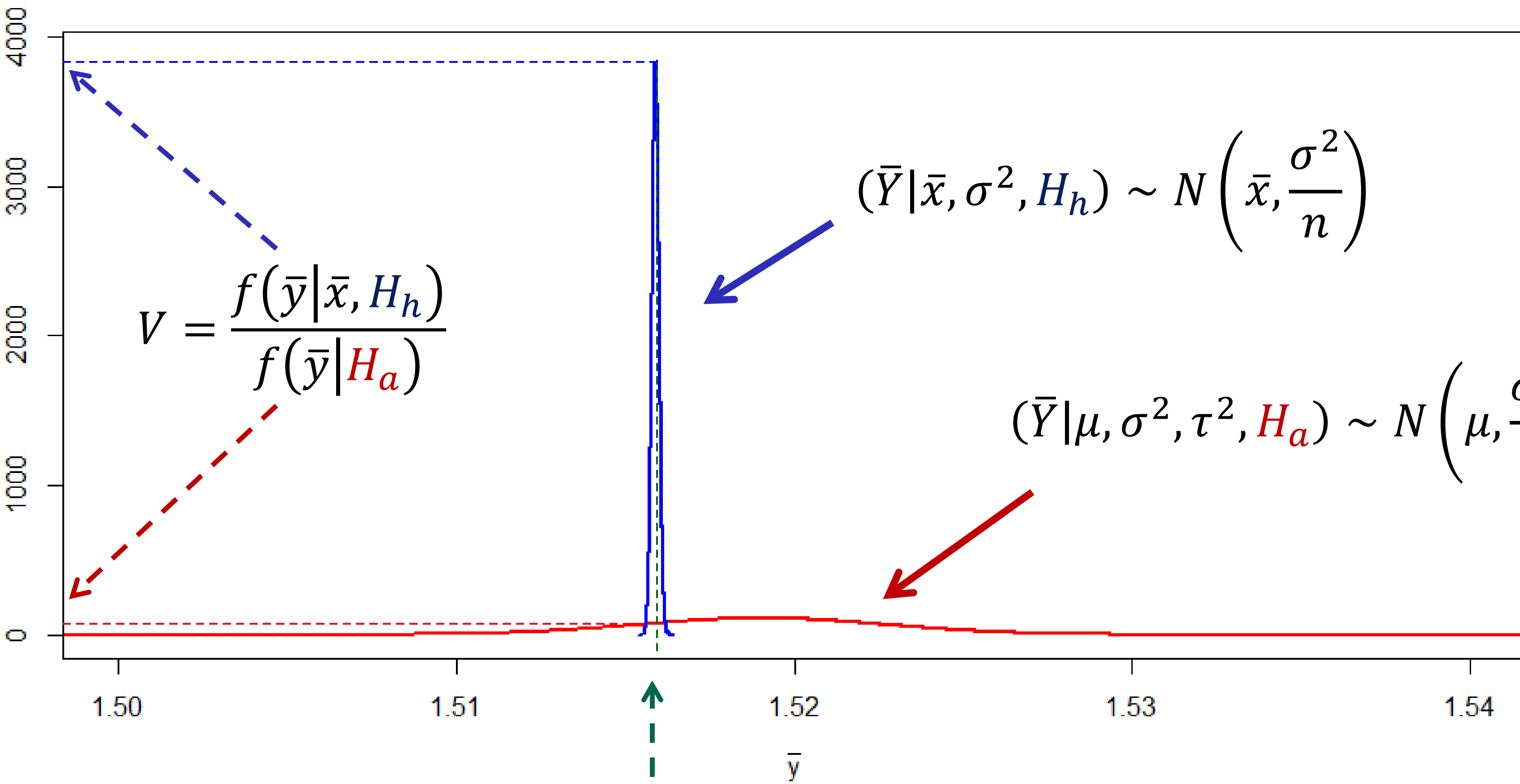
$$V = \frac{f(\mathbf{y} | \mathbf{x}, H_h)}{f(\mathbf{y} | H_a)} = \frac{f(\bar{y} | \bar{x}, H_h)}{f(\bar{y} | H_a)}$$

...och uträkningar ger

$$V = \frac{\sqrt{\sigma^2/n + \tau^2}}{\sqrt{\sigma^2/n}} \times e^{\left\{\frac{(\bar{y}-\mu)^2}{2\cdot(\sigma^2/n+\tau^2)} - \frac{(\bar{y}-\bar{x})^2}{2\cdot\sigma^2/n}\right\}}$$

Polisen

H_h : Glasfragmentet kommer från den krossade rutan
H_a : Glasfragmentet kommer från något annat glasföremål
E: y = (1,51592, 1,51591, 1,51594)
x = (1,51591, 1,51591, 1,51593)



$$V = \frac{f(\bar{y}|\bar{x}, H_h)}{f(\bar{y}|H_a)}$$

$$(\bar{Y}|\bar{x}, \sigma^2, H_h) \sim N\left(\bar{x}, \frac{\sigma^2}{n}\right)$$

$$(\bar{Y}|\mu, \sigma^2, \tau^2, H_a) \sim N\left(\mu, \frac{\sigma^2}{n} + \tau^2\right)$$

$$\bar{y} = 1{,}515923$$

Polisen

$$V = \frac{\sqrt{\sigma^2/n + \tau^2}}{\sqrt{\sigma^2/n}} \times e^{\left\{\frac{(\bar{y}-\mu)^2}{2\cdot(\sigma^2/n+\tau^2)} - \frac{(\bar{y}-\bar{x})^2}{2\cdot\sigma^2/n}\right\}}$$

H_h : Glasfragmentet kommer från den krossade rutan
H_a : Glasfragmentet kommer från något annat glasföremål
E: **y** = (1,51592, 1,51591, 1,51594)
x = (1,51591, 1,51591, 1,51593)

Med exemplets data: $\sigma^2 \approx 1{,}08\cdot 10^{-8}$, $\tau^2 \approx 1{,}003\cdot 10^{-5}$, $\mu \approx 1{,}51909$, $m = n = 3$,
$\bar{y} = 1{,}515923$, $\bar{x} = 1{,}515917$

$$\Rightarrow V = \frac{\sqrt{1{,}08\cdot 10^{-8}/3 + 1{,}003\cdot 10^{-5}}}{\sqrt{1{,}08\cdot 10^{-8}/3}} \times$$
$$\times\, e^{\left\{\frac{(1{,}515923-1{,}51905)^2}{2\cdot(1{,}08\cdot 10^{-8}/3+1{,}003\cdot 10^{-5})} - \frac{(1{,}515923-1{,}515917)^2}{2\cdot 1{,}08\cdot 10^{-8}/3}\right\}} \approx 86{,}6$$

H_h : Glasfragmentet kommer från den krossade rutan
H_a : Glasfragmentet kommer från något annat glasföremål
E: **y** = (1,51592, 1,51591, 1,51594)
x = (1,51591, 1,51591, 1,51593)

B. Exakt variant (Lindley, 1977)

Betingar inte på $\bar{x}$ utan utgår från att alla mätvärden – på omstritt såväl som på jämförelsematerial – ska hanteras genomgående statistiskt.

Uttrycket för V kan utvecklas till:

$$V = \frac{f(\bar{y}|\bar{x}, H_h)}{f(\bar{y}|H_a)} = \cdots = \frac{\int f_{\bar{Y}}(\bar{y}|r,\sigma^2)\cdot f_{\bar{X}}(\bar{x}|r,\sigma^2)\cdot f_{RI}(r|\mu,\tau^2)dr}{\int f_{\bar{Y}}(\bar{y}|r,\sigma^2)\cdot f_{RI}(r|\mu,\tau^2)dr \times \int f_{\bar{X}}(\bar{x}|r,\sigma^2)\cdot f_{RI}(r|\mu,\tau^2)dr}$$

En matematisk hantering!
Transparensen är mycket svår här!

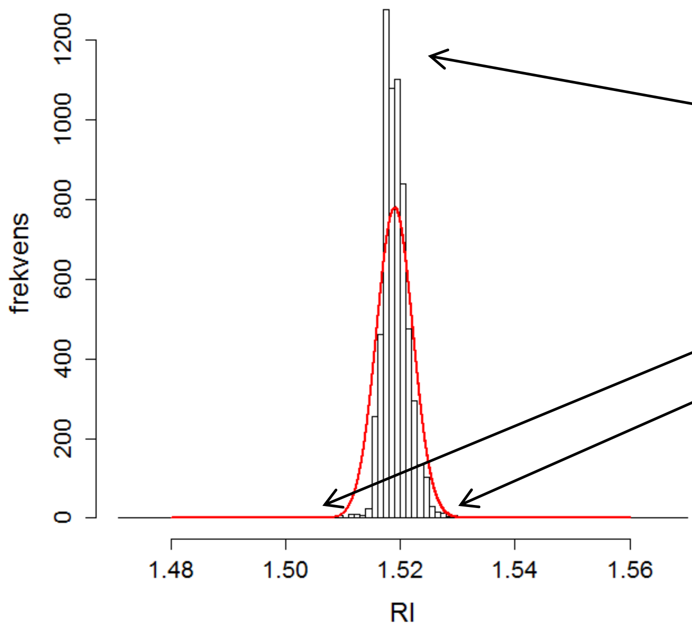
Med exemplets data fås med denna metod $V \approx 61{,}4$

...något svagare grad +1 i skalan

2. Mellanvariationen antas inte normalfördelad

Det är betydligt vanligare att mellanvariationen, dvs. spridningen – i den egenskap som undersöks – mellan objekt med olika ursprung, ***inte*** är normalfördelad

I exemplet med glasdatabasen:



Histogrammet har en spetsigare topp…

… och (kanske) kortare svansar

…än vad den anpassade normalfördelningskurvan har.

Här måste som regel (Lindley's) allmänna uttryck för resultatvärdet användas:

$$V = \frac{\int f_{\bar{Y}}\left(\bar{y}\middle|\theta,\sigma^2\right)\cdot f_{\bar{X}}\left(\bar{x}\middle|\theta,\sigma^2\right)\cdot f_{\Theta}\left(\theta\middle|\mu,\tau^2\right)d\theta}{\int f_{\bar{Y}}\left(\bar{y}\middle|\theta,\sigma^2\right)\cdot f_{\Theta}\left(\theta\middle|\mu,\tau^2\right)d \times \int f_{\bar{X}}\left(\bar{x}\middle|\theta,\sigma^2\right)\cdot f_{\Theta}\left(\theta\middle|\mu,\tau^2\right)d\theta}$$

Θ är en allmän beteckning för den variabel som beskriver den undersökta egenskapens allmänna förekomst. (I exemplet med brytningsindex stod Θ för *RI* .)

Generellt kan vi nästan alltid räkna med att $\bar{x}$ och $\bar{y}$ är normalfördelade. Variationen hos dessa återspeglar slumpmässiga mätfel.

Däremot kan vi mindre ofta räkna med att Θ är normalfördelad.

Normalfördelning

Icke normalfördelning

$$V = \frac{\int f_{\bar{Y}}\left(\bar{y}\middle|\theta,\sigma^2\right)\cdot f_{\bar{X}}\left(\bar{x}\middle|\theta,\sigma^2\right)\cdot f_{\Theta}\left(\theta\middle|\mu,\tau^2\right)d\theta}{\int f_{\bar{Y}}\left(\bar{y}\middle|\theta,\sigma^2\right)\cdot f_{\Theta}\left(\theta\middle|\mu,\tau^2\right)d \times \int f_{\bar{X}}\left(\bar{x}\middle|\theta,\sigma^2\right)\cdot f_{\Theta}\left(\theta\middle|\mu,\tau^2\right)d\theta}$$

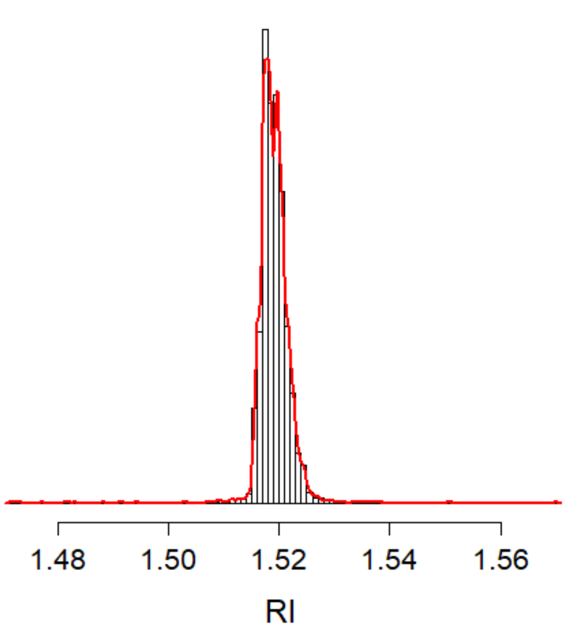
Användande av en annan *parametrisk* fördelning för Θ är svårt – mellanvariation har sällan ett stringent matematiskt mönster $\Rightarrow$ *Kärnskattningsteknik används* (förutsätter mycket data)

Exempel: Brytningsindexen

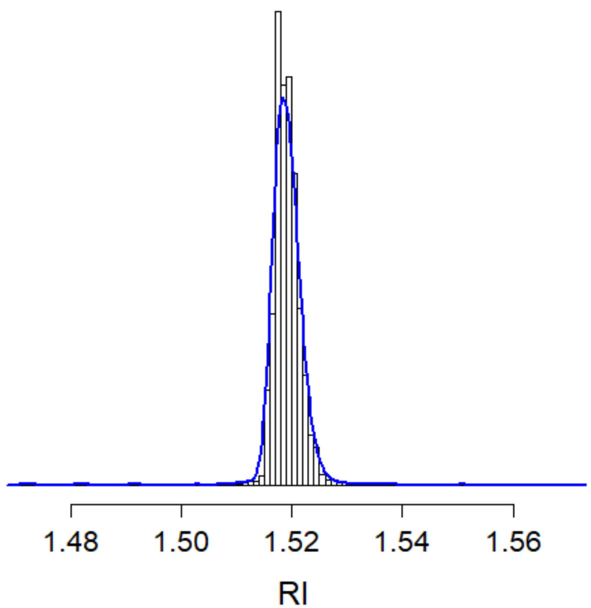
H_h : Glasfragmentet kommer från den krossade rutan
H_a : Glasfragmentet kommer från något annat glasföremål
E: y = (1,51592, 1,51591, 1,51594)
x = (1,51591, 1,51591, 1,51593)

Till glasdatabasen kan vi beräkna en kärnskattning.

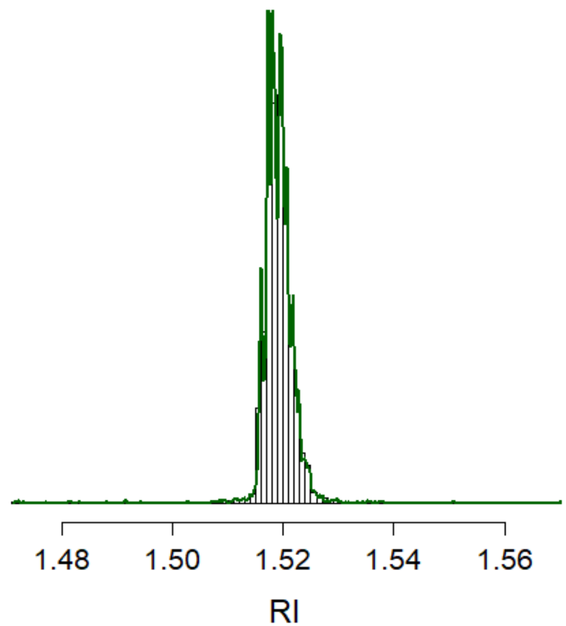
Standardbandbredd (optimal)



Större bandbredd



Mindre bandbredd



$$V = \frac{\int f_{\bar{Y}}\left(\bar{y} \middle| \theta, \sigma^2\right) \cdot f_{\bar{X}}\left(\bar{x} \middle| \theta, \sigma^2\right) \cdot \hat{f}_\Theta\left(\theta \middle| z_1, \ldots, z_k; \lambda\right) d\theta}{\int f_{\bar{Y}}\left(\bar{y} \middle| \theta, \sigma^2\right) \cdot \hat{f}_\Theta\left(\theta \middle| z_1, \ldots, z_k; \lambda\right) d\theta \times \int f_{\bar{X}}\left(\bar{x} \middle| \theta, \sigma^2\right) \cdot \hat{f}_\Theta\left(\theta \middle| z_1, \ldots, z_k; \lambda\right) d\theta}$$

H_h : Glasfragmentet kommer från den krossade rutan
H_a : Glasfragmentet kommer från något annat glasföremål
E: **y** = (1,51592, 1,51591, 1,51594)
x = (1,51591, 1,51591, 1,51593)

Med exemplets data: $\sigma^2 \approx 1{,}08 \cdot 10^{-8}$, $m = n = 3$, $k = 6185$, $z_1, z_2, \ldots, z_{6185}$ tagna ur databasen samt $\bar{y} \approx 1.515923$, $\bar{x} \approx 1.515917$
och med standardbandbredd, som här blir h = 0,0003480376 fås resultatvärdet

$$V \approx 45{,}2$$

Jfr: V beräknat med normalfördelningsantagande:
Evett’s approximativa metod: 86,6
Lindley’s exakta metod: 61,4

Känslighet för val av bandbredd?

Bandbredd	*V*
0,01×standard	45,167
0,5×standard	45,167
2×standard	45,169
10×standard	45,206
100×standard	49,591

Med så pass mycket data i glasdatabasen blir känsligheten låg,

Polisen

Hemuppgift

På en övervakningsbild ser man en person, som relativt oomtvistat är inblandad i en misshandel. I polisutredningen har man anledning att misstänka att en viss Anders Nordgaard är den person man ser på bilden. Vid husrannsakan hos Anders Nordgaard beslagtar man en jacka som man vill jämföra med den jacka som personen på bilden bär.

Jackan tillsammans med övervakningsbilden skickas till NFC för jämförelse. Det kan då sägas vara underförstått att huvudhypotesen i den begärda undersökningen är

H_h : Den beslagtagna jackan är samma jacka som den som bärs av personen på övervakningsbilden

Det framgår ganska tydligt i övervakningsbilden att jackan är svart och av modell CG och den beslagtagna jackan är också en svart CG-jacka.

NFC skriver i sitt utlåtande att den beslagtagna jackan överensstämmer i färg och modell med jackan på bilden, men att ytterligare överensstämmelser inte kunnat konstateras, ej heller några betydelsefulla olikheter. Resultaten talar därför i någon mån för att den beslagtagna jackan och jackan på bilden är densamma (Grad +1).

När utredningen får detta utlåtande från NFC ställer man sig frågande till NFC:s slutsats. Hur kan en positiv slutsatsgrad ha erhållits om man inte sett några andra överensstämmelser än färg och modell. Dessa är väl att betrakta som självklara förutsättningar när jämförelsen ska göras?

Vilken alternativhypotes har man rimligen använt vid NFC för att nå en positiv slutsatsgrad och vilken alternativhypotes tycker man i utredningen borde ha varit den självklara?

Om man ser till hela den forensiska processen, dvs. från misstanke och förodds fram till de efterodds man i utredningen ska uppnå, spelar det någon roll vilken alternativhypotes som används av NFC (förutsatt att man tydligt har angett vilken den är)?