Övning 2 Statistisk översikskurs

Ulf Högnäs

April 2025

Kapitel 11

3. a. Alternative
   b. Null
   c. Alternative
   d. Alternative
   e. Null
   f. Alternative
   g. Null

4. Eftersom *the level of discenrnibility*, eller gränsen för statistisk signifikans, är beräknad med antagandet att nollhypotesen stämmer är svaren 0.05, 0.01 och 0.10.

5. a. Nollhypotesen är att New Yorkers i genomsnitt sover 8 timmar per dag (eller mer). Den alternativa hypotesen är att de i genomsnitt sover mindre än 8 timmar. Vi använder frågeställningens påstående som alternativ hypotes.

   Med matematisk notation: låt μ vara genomsnittlig sömn per natt bland New Yorkers

   $$\begin{aligned} H_0 &: \quad \mu = 8 \\ H_A &: \quad \mu < 8 \end{aligned}$$

   b. Här är nollhypotesen att de anställda inte ägnar mer tid till personligt telefonanvändande (eller e-post) under March Madness. Man jämför med den tid som uppmätts tidigare: i snitt 15 minuter per dag och anställd.

   Nollhypotes: anställda ägnar 15 minuter åt personligt telefonanvändande även under March Madness. Alternativ hypotes: anställda ägnar mer än 15 minuter till detta under March Madness.

   Låt μ vara genomsnittligt användande under M.M.

   $$\begin{aligned} H_0 &: \quad \mu = 15 \\ H_A &: \quad \mu > 15 \end{aligned}$$

7. a. i. Falskt. Vi måste jämföra andelarna.
   ii. Sant. Givet stickprovet så är våra bästa gissningar 3.8% för Rosiglitazone och 3.4% och för Pioglitazone.
   iii. Falskt. Detta är en observationsstudie. Det kan vara något annat förutom medicinen som skiljer sig åt mellan grupperna. De som får Rosiglitazone på recept kan till exempel sämre hälsa i genomsnitt, än den grupp som får Pioglitazone.
   iv. Sant. Eller: vi kan inte säga om skillnaderna beror på medicinen.

   b. $\frac{7979}{227571} \approx 0.035$

   c. $67593 \cdot \frac{7979}{227571} \approx 2370$

   d. i. Om risken för problem och är större med Rosiglitazone.
   ii. Större.
   iii. Den faktiska skillnaden i studien var 0.4%. I diagrammet ser vi att på 100 simuleringar ser vi inte än större skillnad än 0.002 det vill säga 0.2%. Detta motsvarar ett p-värde på minder än 1% (varför?). Vi kan säga att det finns ett samband, men inte vad det beror på.

Kaptiel 12

1. a. Andel i stickprovet som filmats utomhus. Parametern är den verkliga andelen bland alla YouTubevideos och den är okänd.

   b. p och $\hat{p}$

   c. stickprovsandel

   d. Kring stickprovsandelen $\frac{37}{128} \approx 0.289$ w

   e. Svårt att se, men det är 1000 totalt och 900 ska vara i mitten med 50 i varje svans. $(0.22, 0.35)$?

   f. Vi kan säga med 90% konfidens att andelen finns i intervallet $(0.22, 0.35)$

Kaptiel 16

1. (maskinöversatt från boken) För det första ska hypoteserna handla om populationsandelen (p), inte stickprovsandelen.

   För det andra ska nollhypotesens värde vara det vi testar (0.25), inte det observerade värdet (0.29). Det korrekta sättet att formulera hypoteserna är:

$$\begin{aligned} H_0 &: \quad p = 0.25 \\ H_A &: \quad p > 0.25 \end{aligned}$$

3. a. Hypteserna, där p är andelen i Seattle som är för förslaget "defund the police departments"

$$H_0: \quad p = 0.2$$
$$H_A: \quad p > 0.2$$

b.

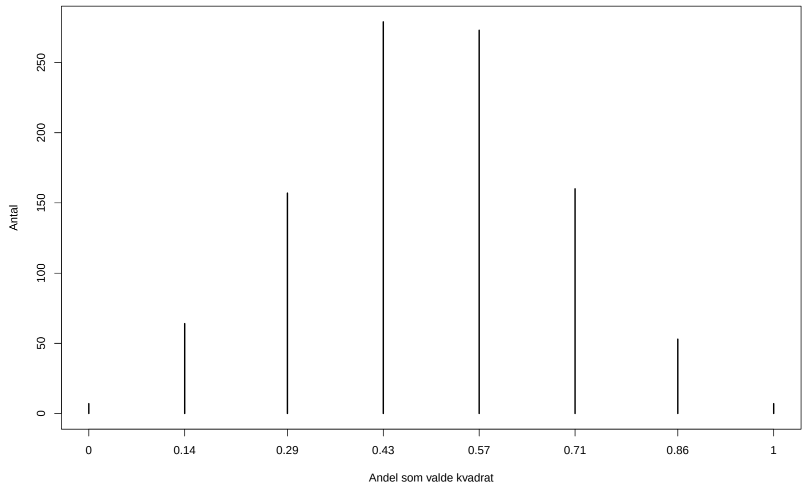
$$\frac{156}{650} \approx 0.2446$$

c. Vi har simulerat med antagandet att den verkliga andelen är 0.2. Det är bara stapeln längst till höger som motsvarar andelar från simuleringen som är större än 0.2446. Det är svårt att se, men denna stapel ska motvara antalet 1. Vi uppskattar därför sannolikheten att få en andel så stor som 0.2446, om den verkliga andelen är 0.20, till en på tusen. Vårt p-värde blir alltså 0.001 eller 0.1%.

Det står inte i problemet vilken gräns för statistisk signifikans som ska användas. Vad tycker du är lagom för en nyhetsartikel av det här slaget?

Låt oss säga att vi, innan undersökningen genomfördes, bestämde oss för gränsen 1%. Eftersom 0.1% är mindre än 1% förkastar vi nollhypotsen. Vi har hittat statistiskt signifikanta bevis för att den verkliga andelen som vill "defund the police departments" är större än 0.2.

Läs även boksens svar!



5. a. Vi ska testa om katter i snitt föredrar den ena formen framför den andra. Vi ska alltså testa om andelen som föredrar kvadraten skiljer sig från 0.5, Låt p vara andelen av alla katter som föredrar kvadraten,

$$\begin{aligned} H_0: &\quad p = 0.5 \\ H_A: &\quad p \neq 0.5 \end{aligned}$$

b. Detta är svårt att avläsa, så jag har gjort en tydligare bild med en tabell (men inte exakt samma simluering, såklart)

Andel	0	0.14	0.29	0.43	0.57	0.71	0.86	1
Frekvens	7	64	157	279	273	160	53	7

Den observerade andelen var $5/7 \approx 0.71$. Vi ser att det sammanlagda antalet simuleringar som fick resultatet 0.71 eller större är

$$160 + 53 + 7 = 220.$$

Vi förväntar oss alltså andelen 0.71 eller större i

$$220/100 = 0.22$$

eller 22% av fallen när nollhypotesen är sann.

Men om två katter hade föredragit kvadraten och fem the Kanizsa square illusion, så hade detta varit lika starka bevis för att katter föredrar den ena över den andra. Vi räknar antalet simuleringar med andelen $2/7 \approx 0.29$ eller mindre

$$7 + 64 + 157 = 228$$

Vi lägger ihop dessa två "svansar" för att få summan av antalet simuleringar som fick ett lika extremt utfall som 5/7, eller mer extremt utfall

$$222 + 228 = 450.$$

Vårt p-värde är därför (approximativt) 0.45. [1].

If I fits I sits: A citizen science investigation into illusory contour susceptibility in domestic cats (Felis silvestris catus)

7. a. I kapitel 16 och föreläsningsanteckningarna kan vi hitta formeln för standardfel (SE) för en stickprovsandel

$$\mathrm{SE} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}.$$

Stickprovsandelen är $\hat{p} = 5/7$ och $n = 7$, så

$$\mathrm{SE} = \sqrt{\frac{\frac{5}{7}(1-\frac{5}{7})}{7}} \approx 0.1707.$$

b. Vi ugår från min tabell igen

Andel	0	0.14	0.29	0.43	0.57	0.71	0.86	1
Frekvens	7	64	157	279	273	160	53	7

Vi kan använda R eller miniräknare tillsammans med tabellen för att räkna ut standardfelet; det blir 0.188. Jag vet inte vad bokens författare har tänkt att man ska göra för att räkna ut standardfelet från diagrammet, men här är min gissning. Vi ordnar alla simuleringar (observationer) i storleksordning. Den 2.5:e percentilen ligger mellan den 25:e och den 26:e observationen, så den 2.5:e percentilen är $\frac{1}{7} \approx 0.14$ (observation 8 till och med 64 är lika med 0.14). På samma sätt kan vi se att den 97.5:e percentilen är $\frac{6}{7} \approx 0.86$. Om andelarna från simuleringen var normalfördelade så skulle de mittersta 95 procenten omfatta allt som är två standardfel från medelvärdet, åt båda hållen. Den totala bredden skulle alltså vara fyra standardfel. Avständet mellan den 2.5:e och den 97.5:e percentilen är $\frac{6}{7} - \frac{1}{7} = \frac{5}{7}$. Så om fyra SE är ungefär 0.72 så måste

$$SE \approx \frac{\frac{5}{7}}{4} = \frac{5}{28} \approx 0.178.$$

Men vi kan inte räkna som om detta är en normalfödelning, se delfråga d. [2]

[1] Det exakta p-värdet kan beräknas i R med `pbinom(2,7,0.5)+(1-pbinom(4,7,0.5))`. Detta ger 0.453125.

[2] Fråga 7 b. är inte en lämplig tentafråga.

c. Ja.

d. Vi kollar om $n \cdot \hat{p} \geq 10$ och om $n \cdot (1 - \hat{p}) \geq 10$.

$$n \cdot \hat{p} = 7 \cdot \frac{5}{7} = 5$$

Detta räcker: kravet är inte uppfyllt och vi kan inte använda sannoliketsmodeller som bygger på den centrala gränsvärdessatsen (och normalfördelningen).

e. Detta är en diskret fördelning med bara åtta möjliga utfall. Normalfördelningen är kontinuerlig.

19. Vi kontrollerar först vår villkor

$$n\hat{p} = 600 \cdot 0.56 = 336 \geq 10$$
$$n(1 - \hat{p}) = 600 \cdot (1 - 0.56) = 264 \geq 10.$$

Båda är större än 10, så stickprovsandelen kommer vara approximativt normalfördelad.

Det betyder i sin tur att om vi skulle kunna upprepa detta stickprov många gånger så skulle ungefär 95% av stickprovsandelarna finnas inom två standardavvikelser från fördelningens mitt.

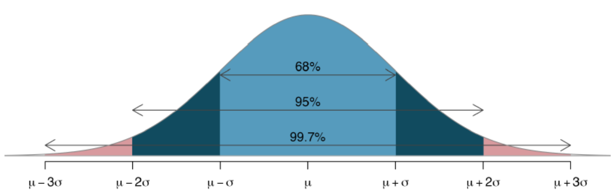


Figure 1: The 68-95-99.7 rule

Vi kommer att använda det mer exakta värdet 1.96, istället för 2. Vi bildar bildar ett intervall kring vår punktskattning som också sträcker sig 1.96 standardavvikelser åt båda hållen. En standardavvikelse som skattas med hjälp av stickprovet kallas Standardfel (SE). Vårt 95%-iga konfidensintervall blir

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

Stickprovsfördelningens mitt är p, den röda linjen i figur 2. Det beror på att stickprovsmedelvärdet $\hat{p}$ är lika med p i genomsnitt. Vi säger att $\hat{p}$ är en *väntevärdesriktig* (engelska *unbiased*) skattning av p.

Figur 2 visar att om det blåa stickprovsmedelvärdet hamnar inom den gråa zonen så fångas det verkliga värdet p. Detta sker med 95% sannolikhet.

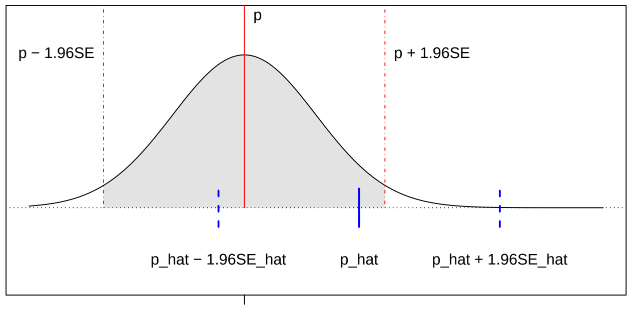


Figure 2: Den verkliga andelen p är okänd

Nu återstår bara att beräkna punktskattningen och standardfelet. Men vi får punkskattningen i problemtexten, $\hat{p} = 0.56$. Formeln för standardfel för en stickprovsandel är

$$\text{SE} = \sqrt{\frac{0.56 \cdot (1 - 0.56)}{600}}$$

Vårt konfidensintervall blir,

$$0.56 \pm 1.96 \cdot \sqrt{\frac{0.56 \cdot (1 - 0.56)}{600}}$$
$$0.56 \pm 1.96 \cdot 0.02026$$

Vi avrundar till detta till

$$(0.520, 0.600)$$

Vi kan säga med 95% konfidens att andelen Kansasbor som planerar att skjuta fyrverkerier den fjärde juli är mellan 52% och 60%.

Sammanfattning, 95%-igt konfidensintervall för en andel

(I) Kontrollera the success-failure condition

(II) beräkna standardfelet (och stickprovsandelen)

(III) använd formeln

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

22. a. Populationsparametern är den verkliga andelen av hela befolkningen som är för legalisering. Den är okänd. Andelen 60% är en sample statistic (stickprovsstatistiska).

b. Vi följer stegen som beskrivs i lösningen till problem 19,

(I) $1563 \cdot 0.60 = 937.8$ och $1563 \cdot (1 - 0.60) = 625.2$

(II)

$$\text{SE} = \sqrt{\frac{0.6 \cdot (1 - 0.6)}{1563}}$$

(III)

$$\begin{aligned}\text{punktskattning} &\pm 1.96 \times \text{SE}\\ 0.6 &\pm 1.96 \cdot \sqrt{\frac{0.6 \cdot (1 - 0.6)}{1563}}\\ 0.6 &\pm 1.96 \cdot 0.01239\\ &(0.575, 0.624)\end{aligned}$$

c. Ja. Syftet med steg (I) är att undersöka om vi har tillräckligt stor stickprovsstorlek i förhållande till skattad andel för att kunna säga att fördelningen är approximativt normalfördelad.

d. Ja. Hela vårt intervall ligger över 50%. Vi kan med 97.5% konfidens säga att andelen som är för legalisering är större än 57.5%. (Varför 97.5% och inte 95%? Rita en bild!)