

F6-8

Statistikens grunder 2 dagtid

HT 2012

Repetition

F6 Uppskattning

Vi skattar populationsparametrar (modellparametrar) med olika statistikor:

Ex. stickprovs-

- -medelvärdet $\bar{X}$ skattar μ
- -variansen S^2 skattar σ^2
- -andelen P skattar π

Kallas för **punktskattningar**.

Grekiskt-romerskt

Ofta men inte alltid:

- Grekiska bokstäver används för parametrar

Ex. $\mu, \sigma^2, \pi, \tau_x$

- Romerska versaler för stokastiska variabler, t.ex. statistikor som skattar parametrar (**extimatorer**)

Ex. $\bar{X}, S^2, P, T_x$

- Romerska gemener för observerade värden

Ex. $\bar{x}, s^2, p, t_x$

Beteckningar, symboler, notation

- Många områden är vana vid "sina" sätt att beteckna
- Det viktigaste är att alltid ta reda på vad som avses och hur de olika beteckningarna definieras.

- Väldigt vanlig beteckning för **estimatorer** är **tak**-symbolen $\wedge$

Ex. μ skattas med $\hat{\mu}$

σ^2 skattas med $\hat{\sigma}^2$

$V(X)$ skattas med $\hat{V}(X)$

Repetition

Osäkerhet i skattningar

En **statistika** som väljs för att skatta en parameter är en s.v.

- Om denna har en **stor varians** så är det större sannolikhet att den ska hamna "långt bort"
 - dvs. **stor osäkerhet**
- Om den har en **liten varians** så är sannolikheten relativt sett mindre att den ska hamna "långt bort"
 - dvs. **liten osäkerhet**

- **Jämför med Figur 16.1 sid 5**

Lite rep

Väntevärdesriktighet 1

- **Eng. unbiased**
- När väntevärdet för estimatoren (dvs. statistikan) är just den parameter vi skattar (estimerar)
- Vi hade sedan förut att $E(\bar{X}) = \mu$ alltså är $\bar{X}$ en **vvr**-skattning för μ
- Hur är det med S^2 är den **vvr**?
- Vi definierade

$$C^2 = \frac{(n-1)}{\sigma^2} S^2 \sim \chi^2(n-1)$$

Väntevärdesriktighet 2

- Vi vet

$$C^2 \sim \chi^2(n-1) \Rightarrow E(C^2) = (n-1)$$

- $E(C^2) = E\left(\frac{n-1}{\sigma^2} S^2\right) = \frac{n-1}{\sigma^2} E(S^2) = (n-1)$

$$\Rightarrow E(S^2) = \frac{(n-1)}{(n-1)} \sigma^2 = \sigma^2$$

alltså är S^2 en **vvr**-skattning för σ^2

Repetition

Konsistens

Lite löst:

- När variansen för estimatoren (dvs. statistikan) går mot noll när $n \rightarrow \infty$
- Osäkerheten minskar ju mer info vi har
- Vi hade redan att $V(\bar{X}) = \sigma^2/n$ alltså är $\bar{X}$ en konsistent skattning för μ
- Hur är det med S^2 ?

Överkurs?

- Vi vet

$$C^2 \sim \chi^2(n-1) \Rightarrow V(C^2) = 2(n-1)$$

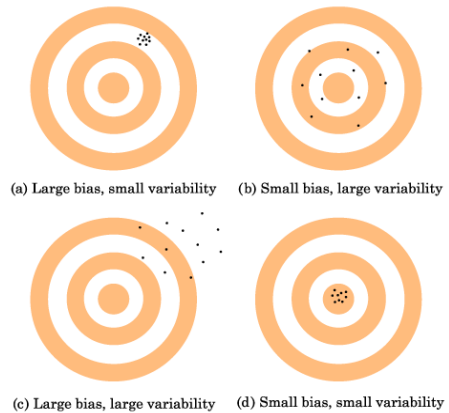
$$\begin{aligned} \bullet \quad V(C^2) &= V\left(\frac{n-1}{\sigma^2} S^2\right) = \left(\frac{n-1}{\sigma^2}\right)^2 V(S^2) = 2(n-1) \\ \Rightarrow V(S^2) &= \frac{2(n-1)}{(n-1)^2} \sigma^4 = \frac{2\sigma^4}{n-1} \end{aligned}$$

alltså är S^2 en **konsistent** estimator för σ^2

Repetition

Skatta μ

- Väntevärdesriktig (eng. *unbiased*) och osäkerhet



Skatta μ

Repetition

- Dra ett stickprov av storlek n från en normalfördelad population med känd varians σ^2
- Vi använder $\bar{X}$ som **punktskattning** för μ
- Vi vill **uttala oss om osäkerheten** kring punktskattningen.

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

Osäkerhetsintervall

Repetition

- Sannolikheten att $\bar{X}$ hamnar i intervallet $\mu \pm 1,96 \cdot \frac{\sigma}{\sqrt{n}}$

$$P\left(\mu - 1,96 \cdot \frac{\sigma}{\sqrt{n}} < \bar{X} < \mu + 1,96 \cdot \frac{\sigma}{\sqrt{n}}\right)$$

$$= P\left(-1,96 < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < 1,96\right) = 0,95$$

- Byt plats på $\bar{X}$ och μ , dvs. låt **gränserna** för intervallet vara **slumpmässiga**

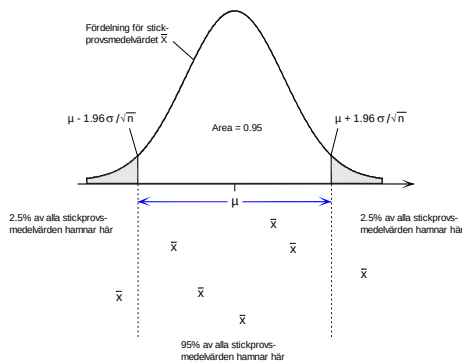
$$P\left(\mu - 1,96 \cdot \frac{\sigma}{\sqrt{n}} < \bar{X} < \mu + 1,96 \cdot \frac{\sigma}{\sqrt{n}}\right)$$

$$= P\left(\bar{X} - 1,96 \cdot \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1,96 \cdot \frac{\sigma}{\sqrt{n}}\right)$$

$$= 0,95$$

Sannolikheten att ligga i ett intervall *Repetition*

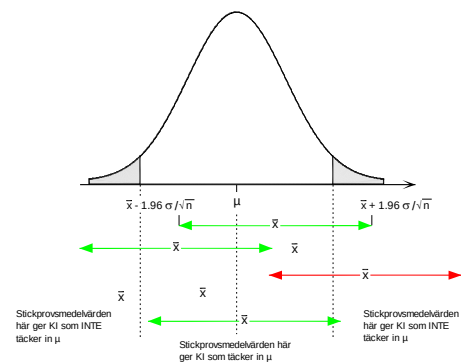
- Upprepade stickprov ger olika medelvärden
- 95 % sannolikhet att $\bar{X}$ hamnar i intervallet



Lägg intervallet runt det observerade $\bar{x}$ istället!

Stokastiska intervall *Repetition*

- 95 % av alla möjliga intervall täcker μ
- Man skattar ett intervall, inte bara en punkt



Jämför med Figur 16.2 sid 8

Exempel och övningar 1

I ett slumpmässigt stickprov med $n = 16$ observationer från en normalfördelning med standardavvikelse $s = 6$ observerades $\bar{x} = 25$.

Skatta μ och ge ett 95 % KI för skattningen

$$n = 16; \bar{x} = 25; z_{0,025} = 1,96$$

$$25 \pm 1,96 \frac{6}{\sqrt{16}}$$

$$25 \pm 2,94 \quad \text{el.} \quad (22,06, 27,94)$$

Exempel och övningar 2

Sex stycken bilar av ett visst märke och årsmodell undersöktes map bensinför-brukningen. Man observerade ett medel-värde = 19,48 och varians = 0,96 (USA-mått!)

Skatta μ och ge ett 95 % KI för skattningen

$$n = 6; \bar{x} = 19,48; s^2 = 0,96$$

$$t_{0,025}^{(5)} = 2,571$$

$$19,48 \pm 2,571 \sqrt{\frac{0,96}{6}}$$

$$19,48 \pm 1,0284 \quad \text{el.} \quad (18,45, 20,51)$$

Skatta en andel 1

- Man vill skatta **andelen** π av en population eller grupp som besitter en viss egenskap:
 - Ska rösta på ett visst parti
 - Bor i villa
 - Vill inte vara med i undersökningen
 - Överlever ett experiment
 - Andelen gånger jag vinner i poker
- Se det också som utfall från en modell eller fördelning:
- Bernoulli-försök
 - lyckas - misslyckas

Skatta en andel 2

- $X_i \sim \text{Bernoulli}(\pi)$, $i = 1, 2, \dots, n$
- $Y =$ summan av X_i
- $Y \sim \text{Bin}(n, \pi)$
- Väntevärde $E(Y) = n\pi$
- Varians $V(Y) = n\pi(1 - \pi)$
- Enligt CGS:

$$Y \rightarrow N(n\pi, n\pi(1 - \pi))$$

$$P = \frac{Y}{n} = \bar{X} \rightarrow N\left(\pi, \frac{\pi(1-\pi)}{n}\right)$$

Men vi vet inte vad π är så vad gör vi när vi ska beräkna variansen som ju behövs till KI?

Formel: K.I. för π

5. $Y \sim \text{Bin}(n, \pi)$, binomialfördelad ($n =$ antal observationer och sannolikhet π att ett objekt har egenskapen vars andel i vill skatta)

$$\frac{y}{n} \pm z_{\alpha/2} \sqrt{\frac{(y/n) \cdot (1 - y/n)}{n}} \quad \text{Tabell 2}$$

$$p \pm z_{\alpha/2} \cdot \sqrt{\frac{p(1-p)}{n}}, \quad p = \frac{y}{n}$$

Använd det näst bästa, dvs. det observerade värdet p

Exempel och övningar 4

$n = 344$ företag tillfrågades om deras policy för gåvor till inköparna från leverantörer. $Y = 83$ företag hade en policy, 261 hade ingen policy.

Skatta andelen företag som har en policy och ge ett 90 % KI för skattningen

$$n = 344; \quad p = \frac{83}{344} = 0,241; \quad z_{0,05} = 1,645$$

$$0,241 \pm 1,645 \sqrt{\frac{0,241 \cdot 0,759}{344}}$$

$$0,241 \pm 0,038 \quad \text{el.} \quad (0,203, 0,279)$$

alt. (20,3 %, 27,9 %)

Ändliga populationer 1

- När man drar med återläggning påverkas inte nästa dragning av tidigare dragningar
- Oberoende**
- När man drar utan återläggning påverkas nästa dragning; om ett objekt redan dragits kan den inte observeras igen
- Beroende**
- Detta påverkar utfallsrummet av möjliga stickprov

Ändliga populationer 2

- Om vi låter $n \rightarrow N$ som är ett ändligt tal och urvalet sker utan återläggning (typiskt) då kommer osäkerheten kring populations-medelvärde minska.
– även andra populationsparametrar
- När $n = N$ finns ju ingen osäkerhet!
- Variansen måste justeras

Ändliga populationer 3

- Varians om $N \rightarrow \infty$ eller om urvalet sker med återläggning:
$$V(\bar{X}) = \frac{\sigma^2}{n}$$
- Varians om N är ändligt och om urvalet sker utan återläggning
$$V(\bar{X}) = \frac{\sigma^2}{n} \cdot \frac{N-n}{N-1} \quad \frac{N-n}{N-1} \rightarrow 0 \text{ när } n \rightarrow N$$
- Ändlig populationskorrektur
eng. finite population correction or fpc

Repetition

Konfidsensintervall 2

Notera att för givet värde på σ^2 så bestäms intervallets längd, dvs. felmarginalen, av

- Konfidsensgraden**
 - 95 % ger oss värdet 1,96 (Tabell 2)
 - Vilket värde skulle vi få vid 90 %?
 - 99 %?
- Stickprovsstorleken**
 - eftersom $V(\bar{X}) = \frac{\sigma^2}{n}$
 - större stickprov $\rightarrow$ mer information

Dimensionera urvalstlk

- Om man vill få bättre skattningar
 - dvs. skattningar med mindre osäkerhet
 - dvs. smalare KI
 vad ska vi göra?
- Större stickprov! Hur stort?
- Nyquist avsnitt 16.5
- Antag att vi ska skapa ett KI t.ex.
 $\hat{\mu} \pm \text{felmarginal}$
 men vi vill att felmarginal $< B$

Dimensionera urvalstlk

- Exempelvis för normalfördelning med känd varians
- Vi vill att $z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq B$
- Lös ut n: $n \geq z_{\alpha/2}^2 \frac{\sigma^2}{B^2}$
- Ex: antag att vi vill ha ett 95 % KI där $\sigma = 10$ och felmarginalen får högst vara ± 2 . Insättning ger

$$n \geq 1,96^2 \frac{10^2}{2^2} = 96,04 \quad \text{Sätt } n = 97$$

Repetition

F7 Sammanfattning

- **Estimatorer** (statistikor, s.v.) skattar parametrar
- En estimator har en fördelning; väntevärde & varians
- Det observerade värdet blir en **punktskattning**
- Osäkerheten formuleras som ett osäkerhetsintervall.
- Punktskattning $\pm$ felmarginal
- **Felmarginalen** bestäms av variansen och **konfidensgraden**

Repetition

Formel: K.I. för π

5. $Y \sim \text{Bin}(n, \pi)$, binomialfördelad (n = antal observationer och sannolikhet π att ett objekt har egenskapen vars andel i vill skatta)

$$\frac{y}{n} \pm z_{\alpha/2} \sqrt{\frac{(y/n) \cdot (1 - y/n)}{n}} \quad \text{CGS} \quad \text{Tabell 2}$$

$$p \pm z_{\alpha/2} \cdot \sqrt{\frac{p(1-p)}{n}}, \quad p = \frac{y}{n}$$

Använd det näst bästa, dvs. det observerade värdet p

Repetition

Ändliga populationer

- Varians om $N \rightarrow \infty$ eller om urvalet sker med återläggning:

$$V(\bar{X}) = \frac{\sigma^2}{n}$$

- Varians om N är ändligt och om urvalet sker utan återläggning

$$V(\bar{X}) = \frac{\sigma^2}{n} \cdot \frac{N-n}{N-1} \quad \frac{N-n}{N-1} \rightarrow 0 \text{ när } n \rightarrow N$$

- Ändlig populationskorrektion
eng. finite population correction or fpc

Repetition

Dimensionera urvalsstk

- Exempelvis för normalfördelning med känd varians

- Vi vill att $z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < B$

- Lös ut n : $n > z_{\alpha/2}^2 \frac{\sigma^2}{B^2}$

- Avrunda högerledet uppåt till närmaste heltal
- Kan även (se Nyqvist) kombineras för fallet med ändlig population

Exempel och övning

Totalt 73 banktjänstemän tillfrågades om de ansåg att det förekom praxis i deras bransch som de ansåg vara oetisk varav 39 svarade "nej".

Skatta andelen som anser att det inte förekommer oetisk verksamhet och ge ett 95 % KI för skattningen.

Lösning:

Först, motivera val av metod.

Kan vi använda CGS?

Är $n = 73$ "tillräckligt stort"?

Tumregel $n > 30$

Exempel och övning, forts

Vi antar att 73 är tillräckligt stort (om inte annat så för att genomföra övningen) och sätter vår tillit till CGS.

Ställ upp ingående storheter, formel som används och utför beräkningen:

$$n = 73; \quad p = \frac{39}{73} = 0,534; \quad z_{0,025} = 1,96$$

$$p \pm z_{\alpha/2} \sqrt{\frac{p \cdot (1-p)}{n}}$$

$$0,534 \pm 1,96 \sqrt{\frac{0,534 \cdot (1-0,534)}{73}}$$

$$0,534 \pm 0,114 \quad \text{el.} \quad (42,0 \% , 64,8 \%)$$

Exempel och övning, forts

Nu fick man (i efterhand) veta att det rörde sig om en enda (liten) bank och att antalet anställda banktjänstemän (som ingår i populationen) är $N = 184$. Vi gör om beräkningen:

*Urvalsfraktionen är
73/184 = ca 40 %*

$$p \pm z_{\alpha/2} \sqrt{\frac{p \cdot (1-p)}{n} \cdot \frac{N-n}{N-1}} \quad \leftarrow \text{fpc = korrektion för ändlig population}$$

$$0,534 \pm 1,96 \sqrt{\frac{0,534 \cdot (1-0,534)}{73} \cdot \frac{184-73}{184-1}}$$

$$0,534 \pm 0,0891 \quad \text{el.} \quad (44,5 \% , 62,3 \%)$$

Att jämföra med (42,0 % , 64,8 %)

Exempel och övning, forts

Nu vill man utöka studien till samtliga banker och banktjänstemän. Vi antar nu att $N \rightarrow \infty$ och använder vanliga formeln. Hur stort stickprov krävs om vi vill ha en felmarginal som är mindre än 0,05? Dvs.

$$z_{\alpha/2} \sqrt{\frac{p \cdot (1-p)}{n}} < B = 0,05$$

Sätt $p = 0,5$ (värsta fallet och inte heller oralistiskt med ledning av den första studien). Insättning ger

$$z_{\alpha/2}^2 \frac{p \cdot (1-p)}{B^2} = 1,96^2 \frac{0,25}{0,05^2} = 384,16 < n$$

Svar: Sätt stickprovsstorleken $n = 385$

Slutsats av övningen?

I exemplet nyss blev punktskattningen för andelen π lika med 53,4 %.

Kan man därmed påstå att **fler än hälften** anser att det inte råder oetisk praxis i bankbranschen?

Notera att KI i första fallet (oändlig pop) blev

$$(42,0 \% , 64,8 \%)$$

Detta intervall täcker 50 %

Kanske slumpen (stickprovsdragningen) orsakade att det blev en liten majoritet?

Hypotesprövning

Att pröva påståenden som denna med statistiska metoder.

Behandlas i Nyqvist Kap 17.

- Formulera ett grundantagande, en s.k. **nollhypotes** som betecknas H_0
- Pröva om data stöder detta antagande genom att jämföra mot en **alternativ hypotes** som betecknas H_1 eller H_A

Popper & falsifierbarhet

Karl Popper och Thurén Kap 17
*Jag har i hela mitt liv endast observerat vita svanar men jag har ändå **inget absolut bevis** för min hypotes att alla svanar är vita. Men, så fort jag observerar en annan sorts svan är hypotesen direkt och absolut motbevisad.*

Popper: man kan aldrig verifiera en hypotes bara **falsifiera** den.

En icke-falsifierad hypotes blir som bäst en **provisorisk sanning**.

Inga empiriska sanningar är säkra!

Exempel

En person påstår att han med förbundna ögon kan avgöra om det blir krona eller klave vid kast med mynt.
 Hur ska vi **testa** detta påstående?

Förslag:

Genomför ett experiment med 12 kast och låt X beteckna rätt antal gissade utfall.

Vi antar vidare att det är lika stor sannolikhet för krona som för klave.

Hur är X fördelad?

Exempel, forts.

12 oberoende kast med sannolikhet p för rätt svar, alltså är $X \sim \text{Bin}(12, p)$.

Om personen bara gissar borde han i snitt bara få hälften rätt, dvs. $p = \frac{1}{2}$

Om $p = 1$ så kommer personen att svara rätt varje gång.

En nollhypotes att börja med är $p = \frac{1}{2}$, dvs. personen i fråga gissar.

Vi är m.a.o. skeptiska, vi vill se tydliga "bevis" för att han klarar detta innan vi överger nollhypotesen!

Exempel, forts.

Under antagandet att $p = \frac{1}{2}$ borde de flesta hamna runt $np = 6$ rätt.

Beslutsregel:

En observation x som är tillräckligt stort, säg $x \geq c$, borde medföra att man **förkastar** grundhypotesen att han bara gissar.

Hur stort ska vi välja c ?

Exempel, forts.

Vi kallar c en **kritisk gräns**.

Observationsområdet där man förkastar nollhypotesen (här $x \geq c$) kallas allmänt det **kritiskt område**.

Beräkna sannolikheterna för att observera $X \geq c$ för några olika c :

c	5	6	7	8
$P(X \geq c)$	0,8062	0,6128	0,3872	0,1938

c	9	10	11	12
$P(X \geq c)$	0,0730	0,0193	0,0032	0,0002

Formulera hypoteser

Två **komplementära** påståenden:

- Antingen accepterar vi den ena
- Eller så förkastar vi den (accepterar motsatsen)

Nyqvists exempel sid 1-7 finns **två** möjligheter: $\mu = 200$ eller $\mu = 220$

Föregående exempel finns **två** möjligheter:

H_0 : Han är inte synsk ($p = 0,5$)

H_A : Han är inte synsk ($p > 0,5$)

Definiera ett test 1

Varje test har en **teststatistika** (testvariabel).

Denna är typiskt baserad på den statistika man använder som **punktskattning** för parametern i fokus, dvs. den **parameter** som man vill testa.

Typiskt vet man hur den är **fördelad under nollhypotesen**.

Uttrycket "under nollhypotesen" betyder att vi utgår ifrån att den är sann.

Definiera ett test 2

Ex. Vi vill testa μ .

- $\bar{X}$ används som punktskattning
- Anta att observationerna är **iid** normalfördelade med känd varians σ^2
- Anta att nollhypotesen är

$$H_0: \mu = \mu_0$$

- **Under nollhypotesen** gäller

$$\bar{X} \sim N(\mu_0, \frac{\sigma^2}{n}) \quad Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$$

Definiera ett test 3

Definitionen av ett *kritiska område* som vi kan beteckna C baseras på:

- teststatistikans fördelning
- mothypotesen ← **Hur?**
- signifikansnivån α ← **Va?**

Observation: $Z = z_{\text{obs}}$. Om z_{obs}

- inte ligger i C , accepteras H_0
- ligger i C , förkastas H_0

Feltyper

Sakernas tillstånd kan vara en av två; antingen är H_0 sann eller falsk.

Vi kan fatta ett av två möjliga beslut; man antingen accepterar el. förkastar H_0

Konsekvens	Situation	
Beslut	H_0 sann	H_0 falsk
Acceptera H_0	Rätt beslut	Typ II fel
Förkasta H_0	Typ I fel	Rätt beslut

Signifikansnivå

Signifikansnivå eller *felrisk*

$$P(\text{Feltyp I}) = P(\text{förkasta } H_0 \mid H_0 \text{ sann}) \\ = \alpha$$

- vill man ska vara liten, nära noll

Testets styrka

$$P(\text{förkasta } H_0 \mid H_0 \text{ falsk}) = 1 - \beta$$

- vill man ska vara stor, nära ett
- $P(\text{Feltyp II}) = \beta$

Signifikans och styrka

Vi betingar på "sanningen", den faktiska situationen:

Sannolikhet	Faktisk situation	
Beslut	H_0 sann	H_0 falsk
Acceptera H_0	$1 - \alpha$	β
Förkasta H_0	α	$1 - \beta$

Lägg märke till att sannolikheterna är betingade på den faktiska situationen

Dubbel- och enkelsidiga

Första varianten av test berör endast **en** enkel parameter, t.ex.

$$H_0: \mu = \mu_0$$

där μ_0 är något tal.

Mothypotesen avgör typ av test:

- **Dubbelsidigt test**

$$H_1: \mu \neq \mu_0$$

- **Enkelsidiga test**

$$H_1: \mu < \mu_0 \quad \text{el.} \quad H_1: \mu > \mu_0$$

Dubbelsidigt test

- Nollhypotes och dubbelsidig mothypotes:

$$H_0: \mu = \mu_0$$

$$H_1: \mu \neq \mu_0$$

Antagande:

Observationerna är iid normalfördelade med känd varians σ^2

Testvariabel är den statistika som skattar μ :

$$\bar{X} \sim N(\mu_0, \frac{\sigma^2}{n})$$

eller normaliserad

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \sim N(0,1)$$

Testvariabelns fördelning

- Nollhypotes och dubbelsidig mothypotes:

$$H_0: \mu = \mu_0$$

$$H_1: \mu \neq \mu_0$$

- Testvariabel är den statistika som skattar μ :

$$\bar{X} \sim N(\mu_0, \frac{\sigma^2}{n})$$

eller oftast standardiserad

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \sim N(0,1)$$

Kritiskt område

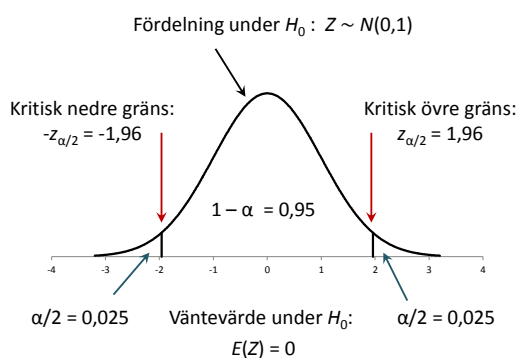
- Givet att H_0 är sann, vad är sannolikheten att observera $\bar{X}$ i intervallet

$$\left(\mu_0 - 1,96 \frac{\sigma}{\sqrt{n}}, \mu_0 + 1,96 \frac{\sigma}{\sqrt{n}} \right)$$

- **Under H_0** är $\bar{X} \sim N(\mu_0, \frac{\sigma^2}{n})$ så sannolikheten är = 0,95.
- Antag att man bestämmer sig för att **förkasta H_0** om man observerar $\bar{X}$ **utanför intervallet**
- Sannolikheten att förkasta H_0 givet att den är sann blir $P(\text{Feltyp I}) = 0,05 = \alpha$, dvs. **signifikansnivån = 5%**

Dubbelsidigt test, forts.

- Om det observerade värdet på Z som vi för tydlighets skull betecknar z_{obs} är tillräckligt långt ifrån nollhypotesens värde μ_0 , förkastas H_0



Repetition

F8 Mer hypotesprövning

Sammanfattning:

- Två komplementära påståenden
- Data ger stöd ("bevis") för en av dem

Vi behöver:

- Nollhypotes och mothypotes
- Signifikansnivå
- Testvariabel och dess fördelning

Ur detta får vi det **kritiska området**.

Repetition

Signifikans och styrka

Signifikansnivå eller felrisk

$$P(\text{Feltyp I}) = P(\text{förkasta } H_0 \mid H_0 \text{ sann}) = \alpha$$

Testets styrka

$$P(\text{förkasta } H_0 \mid H_0 \text{ falsk}) = 1 - \beta$$

$$P(\text{Feltyp II}) = \beta$$

Sannolikhet & konsekvens	Faktisk situation	
	H_0 sann	H_0 falsk
Acceptera H_0	$1 - \alpha$ rätt	β Feltyp II
Förkasta H_0	α Feltyp I	$1 - \beta$ rätt

Arbetsgången 1

1. Förutsättningar / antaganden:

Ett stickprov av storlek $n = 36$ från en normalfördelning med känd varians $\sigma^2 = 9$

2. Nollhypotes: $H_0: \mu = 5$

3. Mothypotes: $H_1: \mu \neq 5$

4. Signifikansnivå / felrisk: $\alpha = 5\%$

5. Teststatistika / testvariabel:

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

Arbetsgången 2

6. Fördelning:

Under de givna förutsättningarna och under H_0 gäller att $Z \sim N(0,1)$.

Om särskilda antaganden behövs, anges dessa, t.ex. "enligt CGS så är Z approximativt $N(0,1)$ ".

7. Kritiskt värde/område:

Vi förkastar H_0 om vi observerar

$$|z_{obs}| > 1.96 = z_{0,025}$$

Detta är ekvivalent med

$$\bar{x} < 5 - 1.96 \cdot \sqrt{\frac{9}{36}} = 4,02$$

$$\bar{x} > 5 + 1.96 \cdot \sqrt{\frac{9}{36}} = 5,98$$

Arbetsgången 3

Alternativ 1.

8. **Beräkningar:** $\bar{x} = 4,01$

9. **Slutsats:** Då $\bar{x} = 4,01 < 4,02$ förkastas H_0 . Det observerade medelvärdet är signifikant skilt från 5 på 5%-nivån. Data stöder inte påståendet att $\mu = 5$.

Alternativ 2.

8. **Beräkningar:** $\bar{x} = 4,03$

9. **Slutsats:** Då $4.02 < \bar{x} < 5.98$ kan H_0 inte förkastas. Det observerade medelvärdet är inte signifikant skilt från 5 på 5%-nivån. Det finns inte tillräckligt stöd i data för att förkasta påståendet att $\mu = 5$.

Medelvärden 1

Fall 1:

1. $X_i \sim N(\mu, \sigma^2)$, σ^2 känd

2. $H_0: \mu = \mu_0$

$$5,6. Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1); \quad \bar{X} \sim N(\mu_0, \sigma^2)$$

7. Kritiskt område:

- enkelsidigt: $z_{obs} > z_\alpha$ ($< -z_\alpha$)

- dubbelsidigt: $|z_{obs}| > z_{\alpha/2}$

Medelvärden 2

Fall 2:

1. $X_i \sim N(\mu, \sigma^2)$, σ^2 okänd

2. $H_0: \mu = \mu_0$

$$5,6. T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t(n-1)$$

7. Kritiskt område:

- enkelsidigt: $t_{obs} > t_\alpha^{(n-1)}$ ($< -t_\alpha^{(n-1)}$)

- dubbelsidigt: $|t_{obs}| > t_{\alpha/2}^{(n-1)}$

Då $n \rightarrow \infty$ går $T \rightarrow Z$. En del författare föreslår att man kan växla till Z om stickprovsstorleken är tillräckligt stor.

Medelvärden 3

Fall 3:

1. $X_i \sim \dots, \sigma^2$ känd, n stort (CGS)
2. $H_0: \mu = \mu_0$

$$5,6. Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \underset{\text{approx}}{\sim} N(0,1); \quad \bar{X} \underset{\text{approx}}{\sim} N(\mu_0, \sigma^2)$$

7. Kritiskt område:
 - enkelsidigt: $z_{obs} > z_\alpha$ ($< -z_\alpha$)
 - dubbelsidigt: $|z_{obs}| > z_{\alpha/2}$

Medelvärden 4

Fall 4:

1. $X_i \sim \dots, \sigma^2$ okänd, n stort (CGS)
2. $H_0: \mu = \mu_0$

$$5,6. Z = \frac{\bar{X} - \mu_0}{S / \sqrt{n}} \underset{\text{approx}}{\sim} N(0,1)$$

7. Kritiskt område:
 - enkelsidigt: $z_{obs} > z_\alpha$ ($< -z_\alpha$)
 - dubbelsidigt: $|z_{obs}| > z_{\alpha/2}$

Egentligen borde testvariabeln betecknas med T efter-som vi använder S istf σ men $T \rightarrow Z$ då $n \rightarrow \infty$. Om man vill vara konservativ använder man t -fördelningen trots att den inte är helt teoretiskt berättigad.

Proportion / andel

Fall 5:

1. $Y \sim \text{Bin}(n, \pi), n$ stort ($Y = n\bar{X} = \sum X_i$)
2. $H_0: \mu = \mu_0$

$$5,6. Z = \frac{Y/n - \pi_0}{\sqrt{\pi_0(1 - \pi_0)/n}} \underset{\text{approx}}{\sim} N(0,1)$$

7. Kritiskt område:
 - enkelsidigt: $z_{obs} > z_\alpha$ ($< -z_\alpha$)
 - dubbelsidigt: $|z_{obs}| > z_{\alpha/2}$

Notera att variansen är känd under H_0 och att man sätter värdet på π till nollhypotesens värde π_0 .

$$Y \underset{\text{approx}}{\sim} N(n\pi_0, n\pi_0(1 - \pi_0))$$

Jämföra grupper 1

- Ofta vill man jämföra två grupper t.ex. med avseende på medelvärden.
- Anta att två grupper A resp. B och deras medelvärden μ_A resp. μ_B ska jämföras
- En typisk hypotesprövning är

$$H_0: \mu_A = \mu_B \Leftrightarrow \mu_A - \mu_B = 0$$
 dvs. ingen skillnad mellan A och B
- Mothypotesen kan vara enkel eller dubbel:

$$H_1: \mu_A > \mu_B \Leftrightarrow \mu_A - \mu_B > 0$$

resp.

$$H_1: \mu_A \neq \mu_B \Leftrightarrow \mu_A - \mu_B \neq 0$$

Jämföra grupper 2

- Man tar ett stickprov från båda grupper, oberoende lika fördelade (*iid*) observationer inom respektive grupp.
- De två **stickproven är oberoende** sinsemellan.
- Detta medför dels att

$$\bar{X}_A \sim N(\mu_A, \sigma_A^2) \quad \bar{X}_B \sim N(\mu_B, \sigma_B^2)$$

och dels att

$$E(\bar{X}_A - \bar{X}_B) = \mu_A - \mu_B$$

$$V(\bar{X}_A - \bar{X}_B) = \frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}$$

$$\bar{X}_A - \bar{X}_B \sim N(\mu_A - \mu_B, \frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B})$$

Gruppjämförelse 1

Fall 6:

$$1. \quad X_{A,i} \sim N(\mu_A, \sigma_A^2), \quad X_{B,i} \sim N(\mu_B, \sigma_B^2)$$

σ_A^2 och σ_B^2 kända

$$2. \quad H_0 : \mu_A - \mu_B = \delta_0 \quad \text{Vanligast är nog } \delta_0 = 0$$

$$5,6. \quad Z = \frac{\bar{X}_A - \bar{X}_B - \delta_0}{\sqrt{\sigma_A^2/n_A + \sigma_B^2/n_B}} \sim N(0,1)$$

7. Kritiskt område:
- se fall 1

Gruppjämförelse 2

Fall 7:

$$1. \quad X_{A,i} \sim \dots, X_{B,i} \sim \dots; \sigma_A^2, \sigma_B^2 \text{ okända}$$

n_A, n_B stora (CGS)

$$2. \quad H_0 : \mu_A - \mu_B = 0$$

$$5,6. \quad Z = \frac{\bar{X}_A - \bar{X}_B - 0}{\sqrt{S_A^2/n_A + S_B^2/n_B}} \underset{\text{approx}}{\sim} N(0,1)$$

7. Kritiskt område:
- se fall 1

Även här kanske beteckna med T istf Z eftersom vi använder S istf σ men $T \rightarrow Z$ då $n \rightarrow \infty$. Om man vill vara konservativ kan man använda t -fördelningen trots att den inte är heller här är helt teoretiskt berättigad.

Gruppjämförelse 3

Fall 8:

$$1. \quad X_{A,i} \sim N(\mu_A, \sigma_A^2), \quad X_{B,i} \sim N(\mu_B, \sigma_B^2)$$

$\sigma_A^2 = \sigma_B^2$ men okända

$$2. \quad H_0 : \mu_A - \mu_B = 0 \quad \text{Viktigt att komma ihåg}$$

$$5,6. \quad T = \frac{\bar{X}_A - \bar{X}_B - 0}{S_p \sqrt{1/n_A + 1/n_B}} \sim t(n_A + n_B - 2)$$

7. Kritiskt område:
- se fall 2

Poolad variansskattning

- Man antar att varianserna är lika i båda grupperna.
- En skattning för den variansen ges av

$$S_p = \sqrt{\frac{(n_A - 1)S_A^2 + (n_B - 1)S_B^2}{n_A + n_B - 2}}$$

Jämföra andelar

Fall 9: $Y_A \sim \text{Bin}(n_A, \pi_A)$, $Y_B \sim \text{Bin}(n_B, \pi_B)$
oberoende stickprov

Nollhypotes: $H_0: \pi_A - \pi_B = 0$

Testvariabel:

$$Z = \frac{\frac{Y_A}{n_A} - \frac{Y_B}{n_B} - 0}{\sqrt{\left(\frac{Y_A + Y_B}{n_A + n_B}\right)\left(1 - \frac{Y_A + Y_B}{n_A + n_B}\right)\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}} \sim N(0,1)$$

Enkelsidigt, dubbelsidigt och kritiska gränser, analogt med tidigare sidor. Kom ihåg att det är CGS och Z-test.

Testa behandlingseffekter

- Ofta vill man jämföra om t.ex. en behandling haft en effekt.
- Man har ett **iid** stickprov av individer som genomgått en behandling säg.
- För **varje individ** i stickprovet finns **två mätningar**:

Före: $X_{i,1}$

Efter: $X_{i,2}$

- De två mätningarna är inte oberoende på individnivå.
- Bilda differenser:

Diff: $D_i = X_{i,1} - X_{i,2}$

D_i 'na är oberoende!

Parvisa differenser

Fall 10:

1. $D_i \sim N(\mu_D, \sigma_D^2)$, σ_D^2 okänd

2. $H_0: \mu_D = 0$ ↗ Det går med andra värden än 0 också

5,6. $T = \frac{\bar{D} - 0}{S_D / \sqrt{n}} \sim t(n-1)$

7. Kritiskt område:
- se fall 2

Då $n \rightarrow \infty$ går $T \rightarrow Z$. En del författare föreslår att man kan växla till Z om stickprovsstorleken är tillräckligt stor.

Pust!

- **Behöver ni komma ihåg alla formler?**
- Nej, de finns med på formelbladet
- **Men vilka behövs eller inte beror på situationen**
- Det som ska läras in är när ni behöver **Z** eller **T** och hur man läser och tolkar formlerna.
- **Kritiska gränser ska ni fastställa själva.**
- Förstå strukturen för test.

p-värden

- Antag att vi får en punktskattning för någon parameter i fokus.
- Antag att vi har en nollhypotes också.
- Men vi har inte bestämt någon signifikansnivå α .
- Antag nu att den kritiska gränsen likställs med punktskattningen.
- **Vad skulle det motsvara för signifikansnivå?**
- Tar vi på måndag ...

Övning 1

Tenta 2010-10-27, uppgift 2.

a) 99 %-igt KI för andelen π

- Låt Y = antalet av 1430 som svarar ja.
- $Y \sim \text{Bin}(1430, \pi)$, binomialfördelad
- $n = 1430$ kan anses tillräckligt stort för att motivera approximering enligt CGS (tumregel $n > 30$).
- Punktskattning $p = 519/1430 = 0,362937$
- Ett 99 %-igt KI för π ges av

$$p \pm z_{0,005} \sqrt{\frac{p \cdot (1-p)}{1430}} \quad (\text{Tabell 2})$$

Övning 1, forts.

- Insättning ger

$$\frac{519}{1430} \pm 2,5758 \sqrt{\frac{519/1430 \cdot 911/1430}{1430}}$$

0,363 $\pm$ 0,033 el. (33,0 % , 39,6 %)

b) Hypotesprövning om π

- Antaganden redan angivna i förra deluppgiften.
- $H_0 : \pi = 0,42$
- $H_1 : \pi < 0,42$
- Signifikansnivå:
Välj själva! Säg $\alpha = 0,05$

Övning 1, forts.

- Testvariabel och dess fördelning:

$$Z = \frac{P - 0,42}{\sqrt{0,42 \cdot 0,58/1430}} \sim N(0,1)$$

- Kritiskt område: förkasta H_0 om

- $Z_{obs} < -1,96 = z_{0,95}$

- $Z_{obs} = \frac{\frac{519}{1430} - 0,42}{\sqrt{0,42 \cdot 0,58/1430}}$
 $= -4,37 < -1,96$

- Slutsats: Vi förkastar H_0 , den observerade punktskattningen är signifikant skild från 0,42 på 5 % nivån.

Övning 2

Tenta 2010-10-27, uppgift 5.

Hypotestest med parvisa differenser;

två avdelningar jämförs där de får samma åtta ärenden och handläggningstid mäts.

- Låt $X_{i,j}$ beteckna handläggningstiden för ärende $i = A, \dots, H$ och avdelning $j = 1, 2$

$$\text{Låt } D_i = X_{i,1} - X_{i,2}$$

Vi antar att $D_i \sim N(\mu_D, \sigma_D^2)$ där σ_D^2 okänd

- $H_0 : \mu_D = 0$
- $H_1 : \mu_D \neq 0$
- Signifikansnivå: $\alpha = 0,05$

Övning 2, forts.

- Testvariabel och dess fördelning:

$$T = \frac{\bar{D}}{S_D/\sqrt{n}} \sim t(7)$$

- Kritiskt område: förkasta H_0 om

$$|t_{obs}| > t_{0,025}(7) = 2,365$$

- Beräkningar:

Ärende	A	B	C	D	E	F	G	H	Σ
Avd 1	10	18	9	27	8	10	8	12	
Avd 2	9	20	9	29	7	12	7	12	
d_i	1	-2	0	-2	1	-2	1	0	-3
d_i^2	1	4	0	4	1	4	1	0	15

Övning 2, forts.

- Data ger $\bar{d} = -3/8 = -0,375$

$$s_d = \sqrt{\frac{\sum d_i^2 - (\sum d_i)^2/n}{n-1}} = 1,407886$$

- Insättning i testvariabeln ger

$$t_{obs} = \frac{-0,375}{1,407886/\sqrt{8}} = -0,75332$$

- Slutsats: H_0 kan inte förkastas, det finns ingen signifikant skillnad på 5 % nivån.