Denne uken: kap. 6.1-6.2-6.3:

Introduksjon til statistisk inferens

- Konfidensintervall
- Hypotesetesting
- P-verdier
- Statistisk signifikans

Slank deg to dager i uken

Skal vekten ned, må kroppen tære på de kiloene den ikke har rukket å forbrenne og derfor spart på til «dårlige tider». Fra tidenes morgen var dette genialt.

I vår tid, og i vår del av verden, er denne kaloriesparingen en nærmest overflødig egenskap. Vi har en nærmest uendelig tilgang på mat av alle slag. Det er ikke greit å være like standhaftig når fristelser lurer overalt.

Nå tyder flere nye undersøkelser på at det kanskje ikke er nødvendig å si nei til litt mer mat hele tiden for å gå effektivt ned i vekt – kanskje tvert imot!

I en nylig publisert britisk studie ble 107 overvektige kvinner (BMI > 30), tilfeldig trukket ut til enten å følge en såkalt kontinuerlig diett på ca. 1500 kcal per dag hele uken gjennom, eller til å leve på en periodevis diett, dvs. ca. 645 kcal i to av ukens syv dager, mens de i ukens øvrige fem dager skulle spise nok kalorier til å vedlikeholde vekten (beregnet individuelt).

Resultatet etter seks måneder var at kvinnene på kontinuerlig diett gikk ned 5,6 kg, mens kvinnene som bare hadde en periodevis kalorireduksjon gikk ned 6,4 kg.

Denne forskjellen var imidlertid ikke statistisk signifikant, og vekttapet må derfor betraktes som likt i de to gruppene.

I begge grupper gikk fastende insulinnivå og grad av insulinresistens noe ned, men mest i gruppen med kalorirestriksjon to dager i uken. Tilsvarende studier viser at periodevis kalorireduksjon også gjør at kroppen bedre tar vare på muskelmassen under vekttapet.

Forskere bak denne typen studier konkluderer med at en periodevis kalorirestriksjon gir like bra vekttap og minst like god helsegevinst som kontinuerlig kalorireduksjon, og dermed kan være et bra alternativ for den som ønsker å gå ned i vekt.

I praksis er dette hyggelige nyheter: Ved å være ekstra oppmerksom på kaloriinntaket et par dager i uken, kan man unne seg litt mer mat og sågar et kakestykke eller et glass vin de øvrige dagene – uten at det går ut over

SØNDAG
Line Kristin Johnson
Klinisk ernæringsfysiolog
vektklubb@vg.no

Resultatet etter seks måneder var at kvinnene på kontinuerlig diett gikk ned 5,6 kg, mens kvinnene som bare hadde en periodevis kalorireduksjon gikk ned 6,4 kg.

Denne forskjellen var imidlertid ikke statistisk signifikant, og vekttapet må derfor betraktes som likt i de to gruppene.

VG 25/9 2011

Statistisk inferens

- Mål: Trekke konklusjoner fra data
- Tidligere har vi undersøkt data og trukket uformelle konklusjoner
- Formell statistisk inferens:
  - Basere konklusjoner på sannsynlighetsberegninger
  - Tar hensyn til usikkerhet/variasjon

Tilfeldig plassering av trær?

- Statistisk analyse:
  - Så mye klustring/sammen-klumping (eller enda mer) vil kun skje i 4% av tilfeller med tilfeldig plassering
  - Observert mønster er altså svært usannsynlig med tilfeldig plassering
  - Konkluderer at det *er* klustring, og ikke tilfeldig plassering

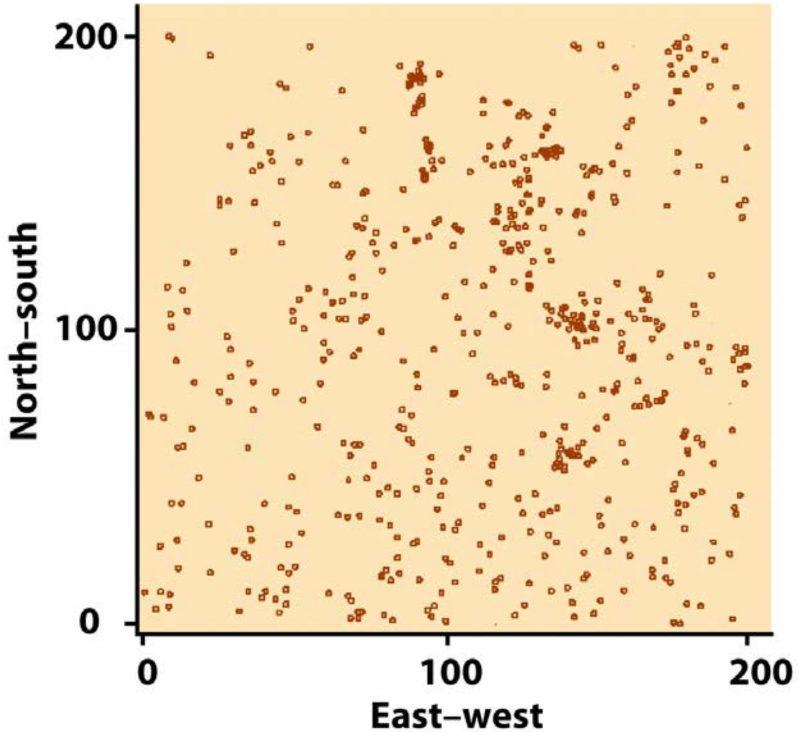

Figure 6-1
Introduction to the Practice of Statistics, Fifth Edition

Effektiv ny medisin?

- Gir ny medisin til 20 pasienter, 12 viser bedring av tilstand
- Gir placebo til 20 andre pasienter, 8 viser bedring av tilstand
- Er den nye medisinen mer effektiv enn placebo?
- Kanskje, men en forskjell som den observerte, eller større, mellom resultatene for de to gruppene vil skje i 20% av slike forsøk bare pga tilfeldig variasjon
- En effekt som så lett kan skje bare ved tilfeldighet er ikke overbevisende nok til å konkludere at den nye medisinen er bedre enn placebo

Formell statistisk inferens

- To viktige metoder
  - *Konfidensintervall*
  - *Signifikanstester*
- Basert på *fordeling* til observator
- Krever *sannsynlighetsmodell* for dataene
- Statistisk inferens baserer seg på at dataene kommer fra et tilfeldig utvalg eller et randomisert eksperiment
  - Viktig å huske på!

Kap. 6: Kjent σ

- Hensikten i dag er å beskrive tankegangen bak statistisk inferens
- Vi skal se på noen enkle metoder for statistisk inferens, som krever en urealistisk antagelse: at vi kjenner det teoretiske standardavviket σ
- Fra og med kap. 7 og utover vil vi se på mer realistiske metoder som kan brukes for de fleste typer av data vi har sett på tidligere

6.1 Konfidensintervall

- SAT-poeng:
  - Dersom poeng for individene i populasjonen er N(μ,σ)-fordelt, vet vi at gjennomsnittet $\bar{x}$ er N(μ,σ/√n)-fordelt
  - Antar at vi vet at σ=100. For n=500 er da σ/√n=4.5
  - 68-95-99.7-regelen: $\bar{x}$ er i [μ-2σ/√n,μ+2σ/√n] = [μ-9,μ+9] med ca 95% sannsynlighet

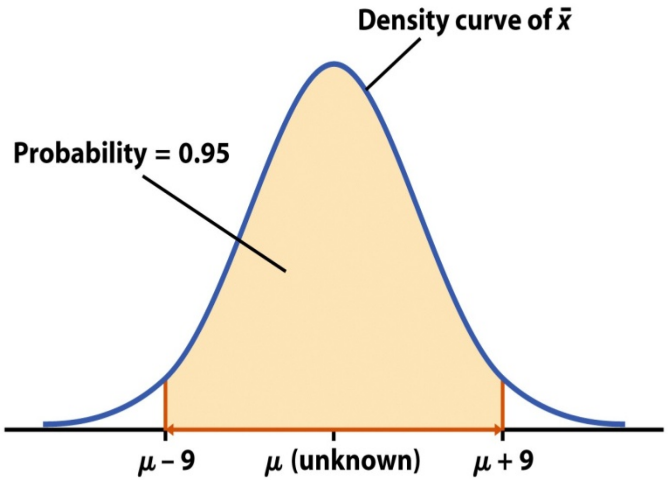

Figure 6-2
Introduction to the Practice of Statistics, Fifth Edition

Konfidensintervall

- SAT-poeng:
  - Dersom peng for individene i populasjonen er $N(\mu,\sigma)$-fordelt, vet vi at gjennomsnittet $\bar{x}$ er $N(\mu,\sigma/\sqrt{n})$-fordelt
  - Antar at vi vet at $\sigma=100$. For $n=500$ er da $\sigma/\sqrt{n}=4.5$
  - 68-95-99.7-regelen: $\bar{x}$ er i $[\mu-2\sigma/\sqrt{n},\mu+2\sigma/\sqrt{n}] = [\mu-9,\mu+9]$ med ca 95% sannsynlighet
  - Å si at $\bar{x}$ er 9 poeng fra μ er det samme som å si at μ er 9 poeng fra $\bar{x}$
  - I 95% av utvalg vil den sanne verdien av μ ligge i intervallet
  - $[\bar{x}-9, \bar{x}+9]$
  - $[\bar{x}-9, \bar{x}+9]$ er et 95% *konfidensintervall* for μ

Konfidensintervall: Tolkning

- SAT-poeng:
  - $[\bar{x}\ -9,\ \bar{x}+9]$ er et 95% *konfidensintervall* for μ
  - Observerer et utvalg med n=500 der $\bar{x}$ =461
  - Vi er 95% sikre (confident) på at den ukjente forventningen μ ligger i intervallet [461-9, 461 +9] = [452, 470]
  - Vi *vet ikke* om vårt utvalg er et av de 95% av utvalgene som har konfidensintervall som fanger μ, *eller* om det er et av de 5% av utvalgene som har konfidensintervall som ikke fanger μ
  - Utsagnet «Vi er 95% sikre (confident) på at den ukjente forventningen μ ligger i intervallet [452, 470]» betyr at vi kom frem til dette intervallet ved å bruke en metode som gir korrekt resultat i 95% av tilfellene

25 utvalg gir 25 forskjellige konfidensintervall, ett dekker ikke μ. Dersom vi tok veldig mange slike utvalg vil 95% av dem gi intervall som dekker den ukjente parameteren μ

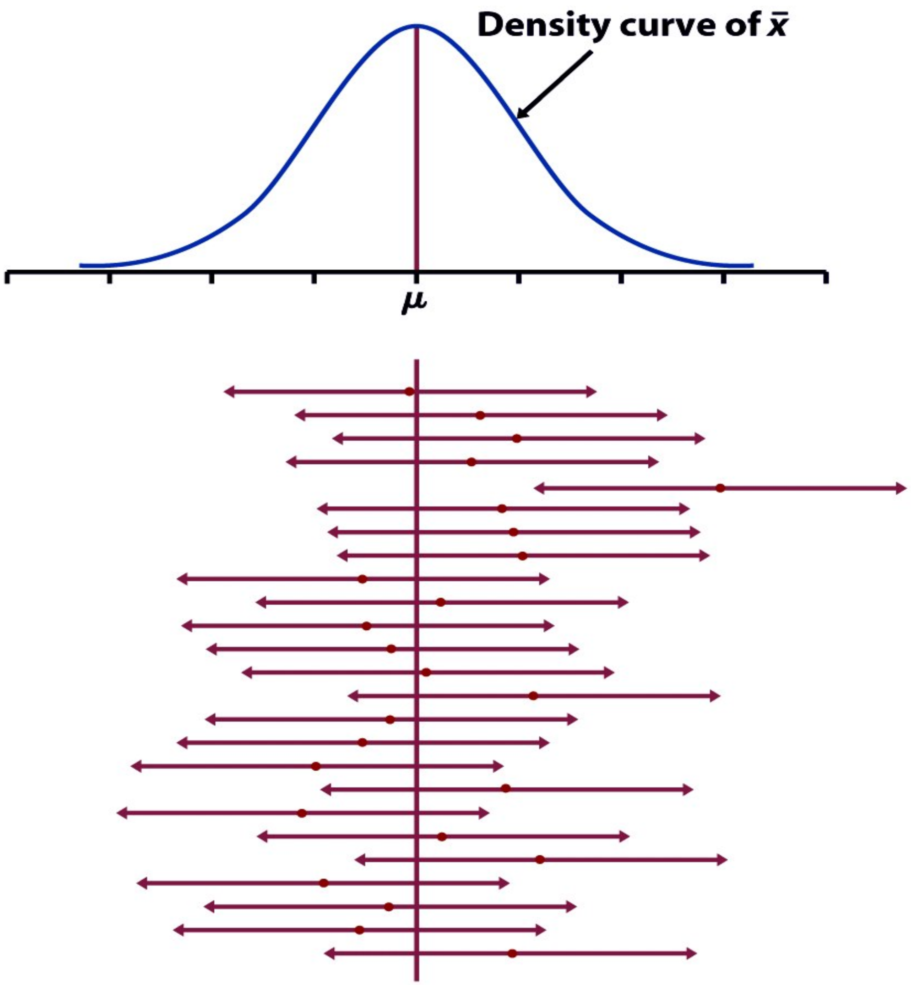

Figure 6-3
Introduction to the Practice of Statistics, Fifth Edition

Konfidensintervall og konfidensnivå

- Man kan velge konfidensnivået, 99, 95 og 90 % er det vanligste. I SAT-eksempelet var 95%-intervallet $\bar{x} \pm 9$
- Generelt konfidensintervall: Estimat $\pm$ feilmargin
- Estimat: Senter i intervall
- Feilmargin: Presisjon av estimat
- Feilmarginen avhenger av konfidensnivå
  - 68-95-99.7-regelen: $\bar{x}$ er i $[\mu-3\sigma, \mu+3\sigma]$ med ca 99.7% sannsynlighet

Konfidensintervall og konfidensnivå

- Generelt konfidensintervall: Estimat $\pm$ feilmargin
- Estimat: Senter i intervall
- Feilmarginen avhenger av konfidensnivå
  - 68-95-99.7 regel: $\bar{x}$ er i [μ-3σ,μ+3σ] med 99.7% sannsynlighet
  - $\bar{x}$ er i [μ-3σ,μ+3σ] er ekvivalent med at μ er i [$\bar{x}$ -3σ, $\bar{x}$+3σ]
  - μ i [$\bar{x}$ -3σ,$\bar{x}$ +3σ] i 99.7% av utvalg
  - [$\bar{x}$ -3σ,$\bar{x}$ +3σ] er et 99.7% *konfidensintervall* for μ
  - Høyere konfidensnivå gir større intervall (større feilmargin)
- *Konfidensnivå* C (95/99.7): Hvor sikre vi er på at konfidensintervallet inneholder sann parameter

Konfidensnivå for forventning

- Normalfordelte data: $\bar{x}$ er eksakt $N(\mu,\sigma/\sqrt{n})$-fordelt
- Sentralgrenseteorem for store utvalg: $\bar{x}$ er tilnærmet $N(\mu,\sigma/\sqrt{n})$-fordelt
- Vi så at vi kunne finne et omtrentlig konfidensintervall for μ ved å bruke 68-95-99.7-regelen
- Skal nå se hvordan vi lager mer presise konfidensintervall for μ
- Må starte med å finne feilmarginene for et bestemt konfidensnivå C
- Går veien om standard normalfordeling: Da kan vi finne generelle feilmarginer som alltid kan brukes for konfidensnivå C når gjennomsnittet $\bar{x}$ er normalfordelt

Konfidensnivå for forventning

- Når Z er N(0,1)-fordelt:
- Velg z* slik at arealet under kurven mellom -z* og z* er

  C, dvs P(-z*<Z <z*)=C
- For tre verdier av C (fra nederste rad i Table D):

z*	1.645	1.960	2.576
C	90%	95%	99%

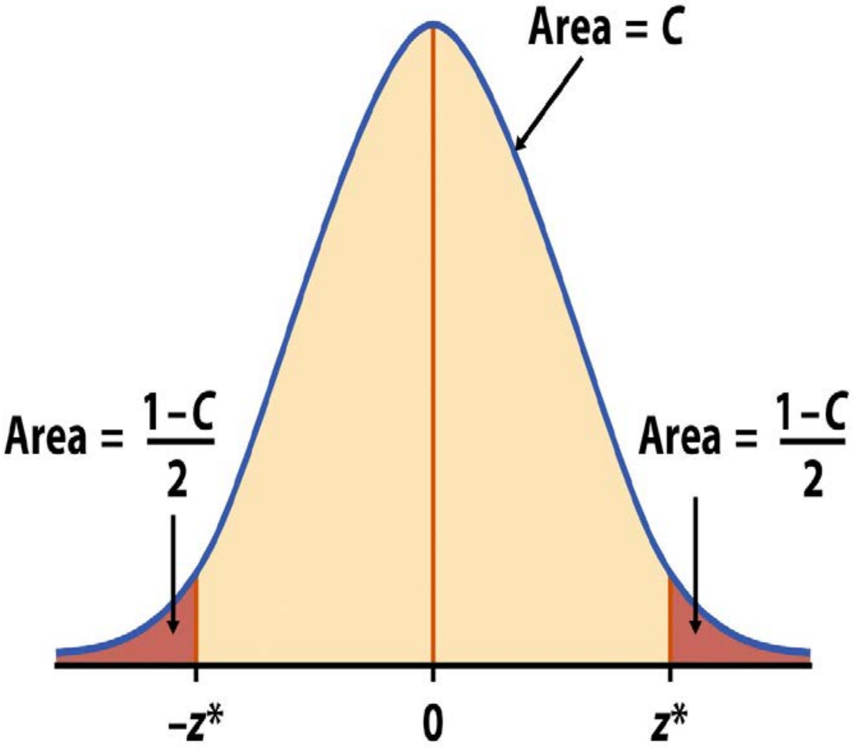

Figure 6-4
Introduction to the Practice of Statistics, Fifth Edition

Konfidensintervall

- $\bar{x}$ er $N(\mu,\sigma/\sqrt{n})$-fordelt, dvs $(\bar{x} - \mu)/(\sigma/\sqrt{n})$ er $N(0,1)$-fordelt
- Fant z^* slik at $P(-z^* < Z < z^*) = C$ når Z er $N(0,1)$-fordelt
- $P(-z^* < (\bar{x}-\mu)/(\sigma/\sqrt{n}) < z^*) = C$
- $P(-z^*\sigma/\sqrt{n} < \bar{x}-\mu < z^*\sigma/\sqrt{n}) = C$
- $P(\bar{x}-z^*\sigma/\sqrt{n} < \mu < \bar{x} + z^*\sigma/\sqrt{n}) = C$
- Konfidensintervall for μ med nivå C: $[\bar{x}-z^*\sigma/\sqrt{n}, \bar{x}+z^*\sigma/\sqrt{n}]$
- Feilmargin $m = z^*\sigma/\sqrt{n}$

CONFIDENCE INTERVAL FOR A POPULATION MEAN

Choose an SRS of size n from a population having unknown mean μ and known standard deviation σ. The **margin of error** for a level C confidence interval for μ is

$$m = z^* \frac{\sigma}{\sqrt{n}}$$

Here z^* is the value on the standard normal curve with area C between the critical points $-z^*$ and z^*. The level C **confidence interval** for μ is

$$\bar{x} \pm m$$

This interval is exact when the population distribution is normal and is approximately correct for large n in other cases.

Definition, pg 388
Introduction to the Practice of Statistics, Fifth Edition

Konfidensintervall: Eksempel

- I en undersøkelse ble et utvalg på 1280 tidligere studenter spurt om hvor stort studielån de hadde
- Gjennomsnittet $\bar{x}$ var \$18900 og standardavviket s var \$49000
- Klart skjev fordeling, men pga stort utvalg er $\bar{x}$ tilnærmet normalfordelt
- Ukjent teoretisk populasjons-standardavvik σ, men later som om det er kjent lik det observerte standardavviket s, dvs $\bar{x}$ er tilnærmet $N(\mu,\sigma/\sqrt{1280})$-fordelt, med σ=\$49.000
- Ønsker å lage et 95%-konfidensintervall for den ukjente forventningen μ
- Vet at z*=1.960 for C=95%

Konfidensintervall: Eksempel

- I en undersøkelse ble et utvalg på 1280 tidligere studenter spurt om hvor stort studielån de hadde
- x er tilnærmet N(μ,σ)-fordelt, med σ=$49000
- Vet at $z^*=1.960$ for C=95%
- Feilmarginen er $m=z^*\sigma/\sqrt{n}=1.960*49000/\sqrt{1280}=2684$, avrunder til m=2700 (fordi $\bar{x}$ er angitt i hele $100)
- 95%-konfidensintervall for den ukjente forventningen μ er $\bar{x}$ +/- m = $18900 +/- $2700 = $[16200, 21600]
- Vi er 95% sikre på at den ukjente forventningen μ ligger mellom $16200 og $21600

Generell form

- Konfidensintervall [$\bar{x}$ -z*σ/√n, $\bar{x}$ +z*σ/√n]

$$\sigma_{\bar{x}} = \sigma / \sqrt{n}$$

$$Konfidensintervall = estimat \pm z * \sigma_{estimat}$$

- *Generell formel* for utvalgsestimater som er normalfordelte (eller tilnærmet normalfordelte)

Hvordan oppfører konfidensintervallene seg?

- Feilmargin $m = z^* \sigma / \sqrt{n}$
- Avhenger av
  - z^* (som avhenger av C): Kan redusere *m* med mindre C
    - Typisk C=95%   (90% og 99% brukes også)
  - σ: Kan redusere *m* med data med mindre variasjon
  - *n*: Kan redusere *m* med flere målinger

Betydning av C (z*)

99% og 95% konfidensintervaller for samme utvalsgsstørrelsen (n=1280)

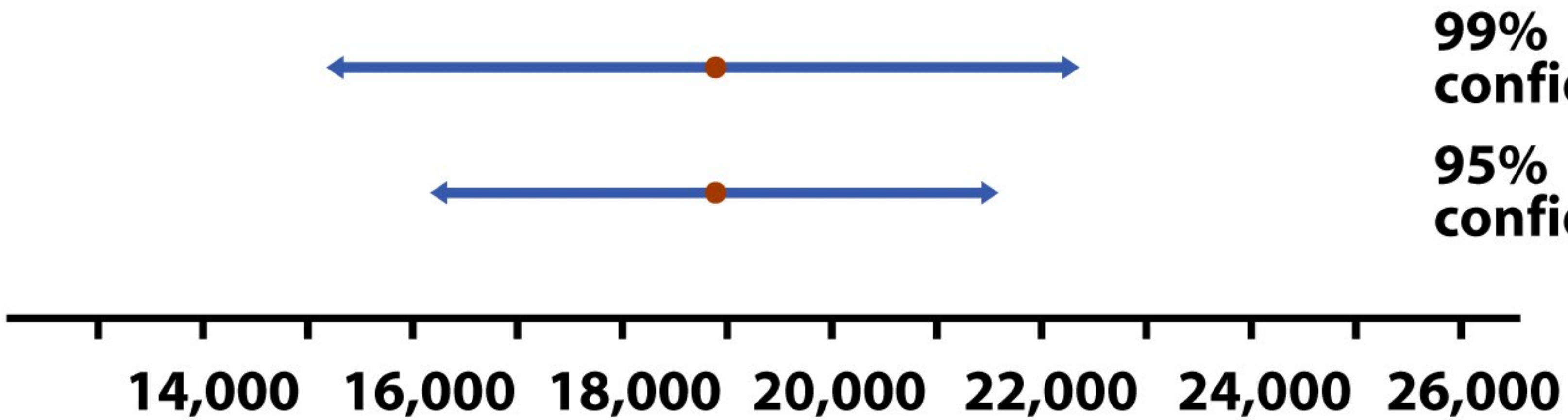

Figure 6-6
Introduction to the Practice of Statistics, Fifth Edition

Betydning av n

95% konfidensintervall for to ulike verdier av utvalsgsstørrelsen n (later som om studielånsundersøkelsen var gjort for et utvalg på n=320 istedet for de virkelige n=1280)

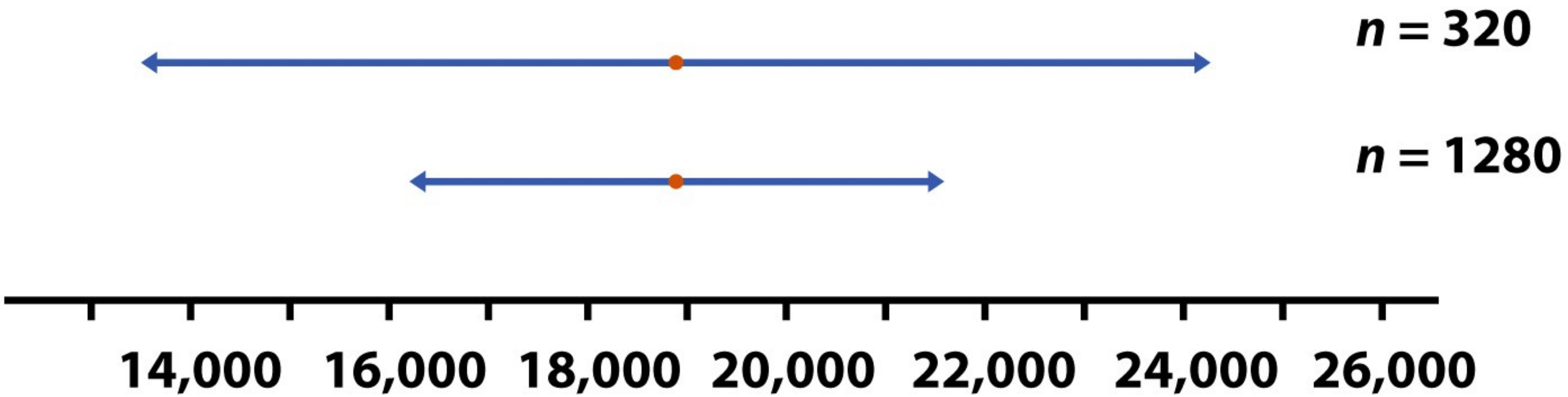

Figure 6-5
Introduction to the Practice of Statistics, Fifth Edition

Valg av utvalgsstørrelse

- Design av eksperiment med bestemt feilmargin *m*:
  - Ønsker feilmargin *m* med konfidens C
  - C gir deg verdien av z*
  - Velg utvalgsstørrelse *n* slik at $m = z^* \sigma / \sqrt{n}$
  - Det vil si velg $n = (z^* \sigma / m)^2$

Eksempel:

Noen forsiktighetsregler

- Data bør være fra et *enkelt randomisert utvalg* (SRS) av populasjonen
  - Viktig med *uavhengige* observasjoner fra populasjonen
- Andre, korrigerte formler for mer kompliserte design
- Følsomt for *uteliggere*
- Lite robust for små n (bør ha n>15) når data ikke er normalfordelte
- Må kjenne σ. Senere skal vi se på hva vi gjør når σ er ukjent
  - Hvis n stor, kan vi bruke $[x - z^*s/\sqrt{n}, x + z^*s/\sqrt{n}]$ (som da er et *tilnærmet* konfidensintervall for μ)