

Oppgave 1 (maks 14 poeng):

1. Forklar hvorfor vi bruker et utvalg fra populasjonen (og ikke hele populasjonen) for statistiske tester og hvordan man gjøre det å trekke et utvalg (angi et eksempel).
2. Formuler en testbar hypotese, og forklar hva type I og type II-feil er. Gi eksempel på de to ut fra hypotesen din. Du kan lage en tabell, men dette er ikke krav.
3. Bruk verdiene nedenfor fra datasettet sykdom.sav, variabel «alder». Beregne standardfeilen og 95% konfidensintervall (se bort ifra at det egentlig er for få datapunkter til å bruke den vanlige konfidensintervallformelen!). Det er valgfri om du vil beregne SE fra hand eller med SPSS. Vis formelen du bruker og rapporter resultatet ditt i APA stil. Verdiene (alder): 46, 44, 45, 43, 75, 24, 28, 52, 36, 93
4. Hva er homoskedastisitet? Forklar kort og gi et eksempel!

Oppgave 2 (maks 16 poeng):

Åpne Athlete.sav og bli kjent med datasettet.

1. Kjør en lineær regresjon med timelønn (salary) som avhengig variabel/utfall og antall medaljer (medal), alder (age), erfaring (years), popularitet i Norge (popularity) og antall timer trening per uke (hours) som prediktorer. Hvilke prediktorer er viktigst? Presenter og tolk resultatet (husk R²) fra analysen. Bruk APA-stil.
2. Forklar med egne ord de viktigste forutsetningene for regresjonen og diskuter om forutsetningene er oppfylt i oppgave 1.
3. Nå vil du vite om det å ha blondt hår eller brunt hår er viktig for suksessen i form av antall medaljer i atletens karriere. Lag en dummyvariabel av disse to hårfargene, og kjør en regresjonsanalyse hvor du undersøker viktigheten av blondt eller brunt hår. Presenter og tolk resultatet (husk R²) fra analysen. Bruk APA-stil. Sjekk forutsetningene her også, men ikke diskuter dem i stor omfang.

Oppgave 3 (maks 24 poeng):

Igen jobber vi med Athlete.sav. Her skal vi gå gjennom alle t-testene vi lærte i forelesningen.

1. Skriv opp formelen til en uavhengig t-test, en avhengig t-test og en one-sample t-test. Beskriv med egne ord hva forskjellen mellom disse typene tester er.

2. Finn ut om det er en signifikant forskjell mellom atletens lønn før (salary) og etter sommersesongen (salary2). Hvilken t-test bruker du her? Er forutsetningene for t-testen oppfylt? Rapportert og tolk resultatene (APA-stil!).
3. Gjennomsnittlig tjener en norsk atlet 31,47 € per time. Er timelønnen i vårt utvalg forskjellig fra dette gjennomsnittet? Er forutsetningene for t-testen oppfylt? Rapportert og tolk resultatene (APA-stil!).
4. a) Er det en kjønnsforskjell mellom menn og kvinner med tanke på timelønn som atlet (salary)? Hvilken t-test bruker du her? Er forutsetningene for t-testen oppfylt? Rapportert og tolk resultatene (APA-stil!). Bruk gjerne en figur for å støtte din tolkning.
b) Er det en forskjell i lønn mellom de med mørk og lys hårfarge? (Lag dummy variabler! For eksempel brun og svart = mørk; blond og rød= lys).
5. Regne ut effekt størrelsen (r eller d) for alle tre tester og diskuter tolkningen av signifikansverdien med tolkningen av effekt størrelsen.

Oppgave 4 (maks 19 poeng):

Bruk samme dataset som før (Athletes.sav).

1. Finn ut om det finnes ulike lønnsnivå i de forskjellige treningssentrene. Rapportert resultatene fra omnibus-testen. Bruk en post hoc-test for å finne ut hvor forskjellen eventuelt ligger. Husk at du må sjekke for lik varianser i alle gruppene. Rapportert og tolk alle resultatene i APA-stil.
2. Finn ut om lønnsnivå er avhengig av både treningscenter og kjønn, og om det er en interaksjon mellom disse to variablene. Husk å rapportere hvor forskjellene ligger innenfor hovedeffektene. Tolk og rapportert resultatene i APA-stil. Sjekk og rapportert forutsetningene her også. OBS: Om varianshomogenitet er ikke gitt i oppgave 2, så antar vi at vår ANOVA er robust nok for å bruke den vanlige F-verdien i dette tilfellet.

PSY2017 STAT KVANT FORSK HØST 2017

EKSAMEN

Generell sensorveiledning

Eksamensoppgaver i høst 2017 bygger på obligatoriske innleveringer studentene har fått i løpet av semesteret. Studentene har vært kjent med oppgavene i lang tid og har jobbet med dem intensivt. Studentene hadde 10 dager etter hver oppgave var publisert på å levere den. Oppgavene ble korrigert og kommentert av assistene. I tilfellet at oppgaven ikke ble godkjent fikk studentene ei uke ekstra på å levere den på nytt. Det var åtte innleveringer til sammen og seks av åtte måtte være godkjente for å få tillatelse til å gå opp til eksamen. Fire oppgaver ble valgt ut til eksamen med små forandringer.

Kursopplegget må tas i betraktning når karakterene settes. Det å ha riktig svar på spørsmålene er grunnforutsetning for å få bestått faget. Forskjellen mellom karakter A og karakter E er derfor graden av presisjon, nøyaktig fremstilling av statistiske analyser og vitenskapelig bruk av språk (korrekt APA-stil!).

Karakter F: Ikke tilfredsstillende minimumskrav, betydelige faglige feil (APA-stil, beregninger, anvendelse, tolkninger av metoder)

Karakter E: Noen faglige feil, med det fleste er riktig løst. Veldig upresist i fremstilling. Feil i APA-stilen.

Karakter D: Alle oppgaver er riktig løst, men fremstilling er svært upresist og ikke i forhold til APA-stilen.

Karakter C: Alle oppgaver er riktig løst, men fremstilling er noe upresist og ikke fullstendig i forhold til APA-stilen.

Karakter B: Alle oppgaver er riktig løst, fremstilling er vitenskapelig presist og i forhold til APA-stilen

Karakter A: Alle oppgaver er riktig løst, fremstilling viser en *fremragende* vitenskapelige presisjon og er fullstendig i forhold til APA-stilen.

Oppgave	Antall poeng	Maks. antall ord
Oppgave 1	14	700
Oppgave 2	16	1000
Oppgave 3	22	800
Oppgave 4	19	900
Total	71	3400 (Tabeller teller ikke med)

Poeng regnes om til karakter etter følgende nøkkel:

- A -> mer enn 62 poeng
- B -> 55-61 poeng
- C -> 48-54 poeng
- D -> 41-47 poeng
- E -> 34-40 poeng
- F -> mindre enn 33 poeng

Oppgave 1 (maks 14 poeng):

5. Forklar hvorfor vi bruker et utvalg fra populasjonen (og ikke hele populasjonen) for statistiske tester, hvordan man trekker ut et utvalg (gi et eksempel) og hva alt det ha å gjøre med signifikans.
6. Formuler en testbar hypotese, og forklar hva type I og type II-feil er. Gi eksempel på de to ut fra hypotesen din. Du kan lage en tabell, men dette er ikke krav.
7. Bruk verdiene nedenfor. Beregne standardfeilen og 95 % konfidensintervall (se bort ifra at det egentlig er for få datapunkter til å bruke den vanlige konfidensintervallformelen!). Det er valgfri om du vil beregne SE for hånd eller med SPSS. Vis formelen du bruker og rapporter resultatet ditt i APA-stil.
Verdiene (alder av deltakere): 46, 44, 45, 43, 75, 24, 28, 52, 36, 93
8. Hva er homoskedastisitet? Forklar kort og gi et eksempel!

Utfyllende sensorveiledning:

Formler (og utregningen) må vises for SE og 95% KI.

- Riktig forklaring av signifikanstesting utvalgstrekning og bakgrunn, sin bruk og utvalgstrekning med et godt eksempel 2.0 poeng
- Riktig forklaring av Type I og Type II feil og god (illustrativ) fremvisning av et eksempel 2.0 poeng
- Riktig beregnet standardfeil, riktig formel 2.0 poeng
- Riktig beregnet 95 % konfidensintervall, riktig formel 2.0 poeng
- Riktig forklaring av homoskedastisitet og god eksempel 2.0 poeng
- Ekstrapoeng avhengig av presisjonsnivået i fremstilling og korrekt bruk av APA stil (0 = riktig men veldig upresist, 4 = fremragende språklige presisjon) Maks. 4.0 poeng

Oppgave 2 (maks 16 poeng):

Åpne Athlete_Eksamen.sav og bli kjent med datasettet.

4. Kjør en lineær regresjon med timelønn (salary) som avhengig variabel og antall medaljer (medal), alder (age), erfaring (years), popularitet i Norge (popularity) og antall timer trening per uke (hours) som prediktorer. Hvilke prediktorer er viktigst? Presenter og tolk resultatet fra analysen.
5. Forklar med egne ord de viktigste forutsetningene for regresjonen og diskuter om forutsetningene er oppfylt i oppgave 1.
6. Nå vil du vite om hårfargen er viktig for suksessen i form av antall medaljer i atletens karriere. Lag dummyvariabler av hårfargene, og kjør en regresjonsanalyse hvor du undersøker viktigheten av hårfargen. (Tips: For de fleste norske atleter er blonde, bruker vi *blond* som baseline). Presenter og tolk resultatet. Sjekk forutsetningene her også, men ikke diskuter dem i stort omfang.

Utfyllende sensorveiledning:

Resultater av regresjon må være rapportert fullstendig og koeffisientene må være riktig betegnet og tolket. Dessuten må alle forutsetninger sjekkes i omfang det er mulig i SPSS.

- | | |
|---|-----------------|
| - Riktig beregnet, fullstendig fremvist og korrekt tolket regresjon i oppgave 1 | 4.0 poeng |
| - Forutsetningene er riktig forklart, beregnet og tolket | 4.0 poeng |
| - Riktig beregnet, fullstendig fremvist og korrekt tolket regresjon i oppgave 3 | 6.0 poeng |
| - Ekstrapoeng avhengig av presisjonsnivået i fremstilling og korrekt bruk av APA-stil (0 = riktig, men veldig upresist, 6 = fremragende språklig presisjon) | Maks. 6.0 poeng |

Oppgave 3 (maks 22 poeng):

Igjen jobber vi med Athlete_Eksamen.sav. Her skal vi gå gjennom alle t-testene vi har lært i forelesningen.

6. Skriv opp formelen til en uavhengig t-test, en avhengig t-test og en one-sample t-test. Beskriv med egne ord hva forskjellen mellom disse typene tester er.
7. Finn ut om det er en signifikant forskjell mellom atletens lønn *før* (salary) og *etter* sommersesongen (salary2). Hvilken t-test bruker du her? Er forutsetningene for t-testen oppfylt? Rapportert og tolk resultatene (APA-stil!).
8. Gjennomsnittlig tjener en norsk atlet 31,47 € per time. Er timelønnen i vårt utvalg forskjellig fra dette gjennomsnittet? Er forutsetningene for t-testen oppfylt? Rapportert og tolk resultatene (APA-stil!).
9. Er det en kjønnsforskjell mellom menn og kvinner med tanke på timelønn som atlet (salary)? Hvilken t-test bruker du her? Er forutsetningene for t-testen oppfylt? Rapportert og tolk resultatene (APA-stil!). Bruk gjerne en figur for å støtte din tolkning.
10. Regne ut effekt størrelsen (r eller d) for alle tre tester fra oppgave 2, 3 og 4 og diskuter tolkningen av signifikansverdien med tolkningen av effekt størrelsen.

Utfyllende sensorveiledning:

- | | |
|--|-----------|
| - Formlene er korrekte og korrekt beskrevet | 3.0 poeng |
| - Riktig t-test ble valgt ut og gjennomført. Forutsetningene er sjekket, resultatene er rapportert i korrekt APA-stil. | 3.0 poeng |
| - Riktig t-test ble valgt ut og gjennomført. Forutsetningene er sjekket, resultatene er rapportert i korrekt APA-stil. | 3.0 poeng |
| - Riktig t-test ble valgt ut og gjennomført. Forutsetningene er sjekket, resultatene er rapportert i korrekt APA-stil. | 4.0 poeng |
| - Korrekt beregnet og tolket effektstørrelser. | 3.0 poeng |

- Ekstrapoeng avhengig av presisjonsnivået i fremstilling og korrekt bruk av APA-stil (0 = riktig men veldig upresist, 5 = fremragende språklig presisjon) Maks. 3.0 poeng
- Ekstrapoeng for gode figurer 3.0 poeng

Oppgave 4 (maks 19 poeng):

Bruk samme datasett som før (Athlete_Eksamen.sav).

3. Finn ut om det finnes ulike lønnsnivå i de forskjellige treningssentrene. Rapportér resultatene fra omnibus-testen. Bruk en post hoc-test for å finne ut hvor forskjellen eventuelt ligger. Husk at du må sjekke for lik varianser i alle grupper og eventuelt bruke robuste tester etterpå. Rapportér og tolk alle resultatene i APA-stil. Gjerne illustrer resultatene dine med en god figur.

4. Finn ut om lønnsnivå er avhengig av både treningssenter og kjønn, og om det er en interaksjon mellom disse to variablene. Husk å rapportere hvor forskjellene ligger innenfor hovedeffektene. Tolk og rapportér resultatene i APA-stil. Sjekk og rapportér forutsetningene her også. Gjerne illustrer resultatene dine med en god figur.

Utfyllende sensorveiledning:

Gjennomsnitt og standardavvik skal rapporteres.

- Del 1: ANOVA, post hoc-testen og test for varianshomogenitet er riktig gjennomført og tolket. 6.0 poeng
- Del 2: ANOVA, post hoc-testen og test for varianshomogenitet er riktig gjennomført og tolket. 6.0 poeng
- Ekstrapoeng avhengig av presisjonsnivået i fremstilling (0 = riktig, men veldig upresist, 5 = fremragende språklig presisjon) Maks. 5.0 poeng
- Ekstrapoeng for gode figurer 2.0 poeng

