

**STUDIEÅRET 2021/2022**

**UTSATT INDIVIDUELL SKRIFTLIG HJEMMEEKSAMEN**

I

**FORSKNINGSMETODE (MET 200)**

Utlevering av oppgave: mandag 14. februar 2022 kl. 10.00

Innlevering av oppgave: mandag 14. februar 2022 innen kl. 14.00 i WISEflow

**NB!** Det skal kun lastes opp én fil som MÅ være i pdf-format.

**Eksamensform og vurdering:**

Individuell, skriftlig hjemmeeksamen (4 timer).

Eksamen består av 3 deler. Hver del har flere deloppgaver som vektes i henhold til poeng oppgitt for hver deloppgave. Eksamen vurderes med gradert karakter (A-F).

Hjelpemidler: Alle tilgjengelige hjelpemidler er tillatt. Deler av eksamen krever utregninger i Excel. Det kreves ikke kildehenvisninger.

## **Kvalitativ metode** (maksimalt 15 poeng)

### **Oppgave 1.** (5 poeng)

Hva er semi-strukturerte intervjuer, og hvorfor er dette en vanlig metode å benytte i kvalitativ forskning?

### **Oppgave 2.** (5 poeng)

Forklar prinsippet om konfidensialitet i kvalitativ forskning, og vis hvordan dette prinsippet kan ivaretas gjennom forskningens ulike faser.

### **Oppgave 3.** (5 poeng).

Redegjør for begrepet forforståelse. Kom med et (tenkt) eksempel på en situasjon hvor forskeren bør reflektere rundt sin egen forforståelse av fenomenet som studeres.

## **Kvantitativ metode** (maksimalt 15 poeng)

### **Oppgave 1.** (5 poeng)

Du får i oppdrag å gjennomføre en studie for å undersøke om det er en sammenheng mellom motorisk utvikling i 4-6 års alder og fysisk aktivitetsnivå som ung voksen. Ta utgangspunkt i at du har tilgang på et representativt utvalg barnehagebarn og at det finnes valide målemetoder til formålet og at det er etisk forsvarlig å gjennomføre studien. Beskriv hvordan du ville designet en slik studie og hvorfor designet du har valgt bringer oss nærmest mulig et svar på om det er en sammenheng mellom motorisk utvikling og aktivitetsnivå som voksen. Sett opp et tosidig hypotesepar for studien. Beskrivelsen bør ikke overskride ½ side.

### **Oppgave 2.** (6 poeng)

Under finner du beskrivelser av tre studier. Avgjør hvilket studiedesign de ulike studiene benytter. Begrunn kort (1-2 setninger).

- a) En gruppe forskere benytter data fra en stor nasjonal kartlegging av fysisk aktivitetsnivå blant barn og unge i Norge. I studien ble det inkludert barn i 1., 4. og 10. klasse fra skoler i hele landet. Alle barna har gått med aktivitetsmåler i 7 dager og ved bruk av spørreskjema har de svart (JA/NEI) på om de har diagnostisert astma. Forskerne ønsker å undersøke om det er en assosiasjon mellom fysisk aktivitet og astma.
- b) En gruppe forskere kobler, ved hjelp personnummer, aktivitetsnivå registrert med akselerometer blant norske menn og kvinner i alderen 20-85 år mot data fra dødsårsaksregisteret. Aktivitetsnivået er registrert for noen år siden. Dette gjør de for å undersøke om det er en sammenheng mellom fysisk aktivitetsnivå og total dødelighet.
- c) En gruppe forskere ønsker å undersøke om 12 uker med henholdsvis intervalltrening og kontinuerlig utholdenhetstrening påvirker oksygenopptaket hos en gruppe utrente menn og kvinner. Forskerne rekrutterer deltakere inn i studien fortløpende og etter at de har samtykket blir tilfeldig fordelt mellom følgende grupper; 1) intervalltrening to ganger i uka, 2) kontinuerlig utholdenhetstrening 2 ganger i uka, eller 3) fortsette som normalt i 12 uker. Målsettingen i studien er å rekruttere 15 deltagere til hver gruppe. Deltagere måler oksygenopptaket før og etter 12 ukers perioden og endringen sammenlignes mellom de tre gruppene.

### **Oppgave 3.** (2 poeng)

Hva innebærer det at en målemetode er valid? Gi et tenkt eksempel på hvordan du kan gå frem for å undersøke om en målemetode er valid.

#### Oppgave 4. (2 poeng)

Hva mener vi når vi sier at et forskningsresultat er generaliserbart? Hvilke forutsetninger ligger til grunn for at vi kan generalisere et forskningsresultat?

#### Statistikk (maksimalt 30 poeng)

Statistikkdelen av eksamen er delt inn i tre deler. Hver del har et tilhørende datasett. De tre datasettene laster dere ned fra Canvas (Moduler → Eksamen - datasett).

#### DEL 1 (Bruk datasettet «datasett\_basis2021\_eksamen.xlsx» i Canvas)

#### Oppgave 1. (4 poeng)

Fyll ut tabellen med passende deskriptiv statistikk for de ulike variablene, separat for menn og kvinner. Rund av til to desimaler der nødvendig. Forklar hvordan du har tenkt.

Variabel	Menn (n=107)	Kvinner (n=93)
LDL-kolesterol (mmol/L)		
MVPA pr dag (min)		
Røykestatus		
Aldri		
Tidligere		
Nåværende		

#### Oppgave 2. (2 poeng)

Hva er forskjellen på gjennomsnitt og median? Vurder hvordan disse sentralmålene påvirkes av ekstreme verdier (ikke behov for Excel).

#### Oppgave 3. (2 poeng)

Når man konkluderer etter en statistisk test er det to typiske feil man kan gjøre når det kommer til hypotesene. Forklar hva feilene kalles og hva hver av disse innebærer (ikke behov for Excel).

#### Oppgave 4. (2 poeng)

Hva er forutsetningene for å kunne gjøre en lineær regresjon (ikke behov for Excel)?

## **DEL 2 (Bruk datasettet «Vekt.xlsx» i Canvas)**

Vi har gjennomført en 12 ukers treningsintervensjon blant overvektige voksne, og vi ønsker å se på hvordan kroppsvekten til deltakerne har endret seg fra før intervensjon (Vekt\_pre) til etter intervensjon (Vekt\_post). Bruk tosidig p-verdi, med signifikansnivå satt til 0,05.

### **Oppgave 1.** (1 poeng)

Hva slags variabeltyper er de sentrale variablene her?

### **Oppgave 2.** (1 poeng)

Hva er passende statistisk test for å svare på problemstillingen?

### **Oppgave 3.** (4 poeng)

Gjør analysen og konkluder på problemstillingen.

### **Oppgave 4.** (2 poeng)

Er forutsetningene for å gjøre den statistiske testen møtt? Utdyp.

## **DEL 3 (Bruk datasettet «Forsikringsutbetalinger.xlsx» i Canvas)**

Du har blitt leid inn som konsulent i et amerikansk helseforsikringsselskap, der de ønsker å finne ut av om de har større årlige dollarutbetalinger («Kostnad») til de som røyker enn de som ikke røyker («Røyker»). Gjør om variabelen «Kostnad» til en kategorisk variabel («Kostnad\_kategori»), der du får oppgitt at kostnader under 10 000 dollar/år klassifiseres som «Lav» utbetaling, kostnader fra 10 000 til 29 999 dollar/år klassifiseres som «Middels» utbetaling, og kostnader på 30 000 dollar/år eller mer klassifiseres som «Høy» utbetaling. Du velger å sette signifikansnivået til 0,05.

### **Oppgave 1.** (1 poeng)

Sett opp  $H_0$  og  $H_1$  for denne problemstillingen.

### **Oppgave 2.** (1 poeng)

Hva slags variabeltype er «Røyker»?

### **Oppgave 3.** (4 poeng)

Lag to krysstabeller med variablene «Røyker» og «Kostnad\_kategori» - en med de observerte verdiene og en med de forventede verdiene.

### **Oppgave 4.** (4 poeng)

Bruk en egnet figur til å visualisere forskjellene mellom disse to variablene. Kalkuler også en p-verdi ved å bruke kji-kvadrat, og konkluder på problemstillingen.

### **Oppgave 5.** (2 poeng)

Er forutsetningene for å kunne gjøre kji-kvadrat møtt? Forklar.