

Ordinær eksamen MET200 – høsten 2022 (maksimalt 60 poeng)

Del 1 – Kvalitativ metode (maksimalt 15 poeng)

1. Nevn minst tre eksempler på kvalitative forskningsmetoder. (3 poeng)

Sensor:

Individuelle intervju, fokusgruppeintervju, observasjon, tekstanalyse, bildeanalyse

2. Redegjør kort for hvordan forskeren sikrer at en kvalitativ studie fremstår troverdig og gyldig. (6 poeng)

Sensor:

Troverdighet: Konkrete og spesifikke beskrivelser av fremgangsmåter i forskningsprosessen. Gjøre forskningsprosessen gjennomiktig. Detaljerte beskrivelser av forskningsstrategi og analysemetoder, forskningsprosessen skal kunne vurderes av utenforstående trinn for trinn.

Gyldighet: Teoretisk gjennomsiktighet. Beskrive grunnlaget for tolkninger og vise hvordan analysen gir grunnlag for konklusjoner og tolkninger. Sikre at forskerens tolkninger er gyldige i forhold til den virkeligheten de beskriver. Deltakernes vurdering, forskerens tilknytning til og kjennskap til miljøet, resultater i forhold til andre studier.

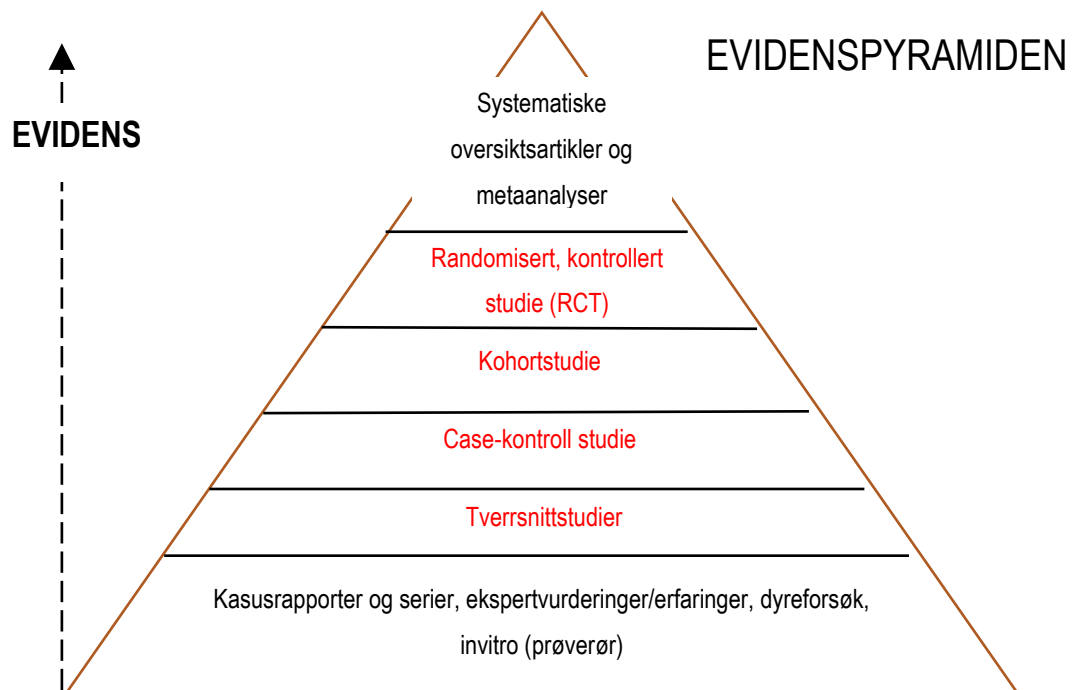
3. Redegjør kort for hva teori kan bidra med i en kvalitativ studie. (6 poeng)

Sensor:

Teori bidrar med forståelse av eksempelvis hvordan samfunn fungerer, hvordan organisasjoner opererer, hvordan og med hvilke begrunnelser personer samhandler på bestemte måter eller hvordan personers erfaringer kan forstås

Del 2 – Kvantitativ metode (maksimalt 15 poeng)

1. Nevn noen sentrale moment som kjennetegner studier gjennomført med kvantitativ metode. (4 poeng)
 - a. Målbare/kvantifiserbare data, større grupper (utvalg fra populasjon), ønske om å generalisere, interessert i det gjennomsnittlige og ikke enkeltpunkter, objektiv forskerrolle (så langt det lar seg gjøre).
2. Fyll inn, på riktig plass, de manglende studiedesignene i evidenspyramiden. Hvis dere ikke får til å fylle inn i pyramiden, kan dere skrive studiedesignene i tekst og i hvilken rekkefølge de er listet (topp til bunn eller bunn til topp). (2 poeng)



3. Hvilke kriterier må (minst) være oppfylt for å kunne kalle en relasjon mellom to variabler for kausal? (3 poeng)
 - a. (1) Tidsrelasjon mellom variablene (X må inntreffe før Y), (2) observert sammenheng mellom X og Y, (3) fravær av andre mulige forklaringer (konfunderende variabler).
4. Forklar hvordan du ville designet en studie som hadde som mål å påvise kausalitet. (4 poeng)
 - a. Vi er ute etter en beskrivelse av sentrale moment ved en randomisert, kontrollert studie (RCT), som er «gullstandard» blant enkeltdesign for å påvise kausalitet. For full pott må momenter som inklusjons- og eksklusjonskriterier,

baselinetest, randomisering til intervensjon eller kontroll, intervensjonsperiode og post-test være del av svaret og i riktig rekkefølge.

5. Hva menes med at en målemetode er valid? Gi et tenkt eksempel. (2 poeng)

- a. En målemetodes validitet omtales også ofte som en målemetode gyldighet, og refererer til i hvilken man nøyaktig klarer å måle det man faktisk ønsker å undersøke. Reliabilitet er en forutsetning for validitet, da målingen ikke kan være valid/gyldig uten å være reproducerbar. F.eks. hvis vi ønsker å måle aerob kapasitet hos deltakerne våre, må vi vite at denne målingen både vil være stabil om den gjøres flere ganger og at den faktisk gir et nøyaktig bilde av den aerobe kapasiteten (direkte måling av VO₂maks) for å være valid.

Del 3 – Statistikk (maksimalt 30 poeng)

Generell del og deskriptiv statistikk

1. Hvilke ulike variabeltyper har vi, og hva slags deskriptiv statistikk bør brukes for å beskrive disse variabeltypene i en tabell? (2 poeng)

a. Vi har i hovedsak kategoriske (nominale og ordinale) og numeriske (diskrete og kontinuerlige) variabeltyper. Numeriske variabler beskrives med gjennomsnitt og standardavvik hvis variabelen er tilnærmet normalfordelt, og med median og kvartiler hvis den er skjevfordelt. Kategoriske data beskrives i hovedsak med antall og prosent (kan evt. legge til at ordinale variabler med 5 eller flere kategorier kan beskrives med median og kvartiler, men ikke nødvendig).

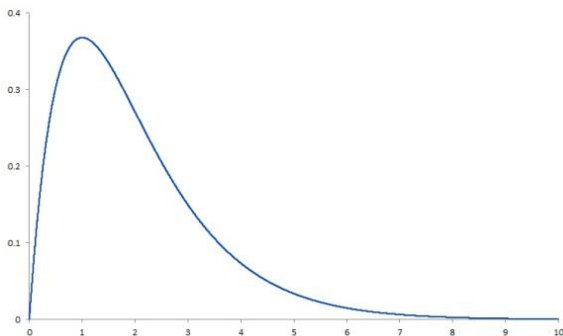
2. Se på tallrekken under. Hva er gjennomsnitt og median i denne tallrekken? Er gjennomsnitt og median her likt/ulikt? Hva påvirker størrelsesforholdet mellom gjennomsnitt og median? (2 poeng)

- a. Gjennomsnitt: 184
b. Median: 147
c. Gjennomsnittet er her høyere enn median. Forholdet mellom gjennomsnitt og median påvirkes av hvor skjevfordelt den numeriske variabelen er, hvor uteliggere/ekstreme verdier vil dra gjennomsnittet i denne retningen mens median er mer robust i møte med ekstreme verdier. I en normalfordelt variabel vil gjennomsnitt og median være likt.

182	89	47	299	411	113	147
-----	----	----	-----	-----	-----	-----

3. Hva kalles sannsynlighetsfordelingen i figuren under? (1 poeng)

- a. Høyreskjev sannsynlighetsfordeling («halen» går ut mot høyre).



4. Definer først begrepene p-verdi og signifikansnivå. Hvordan brukes disse begrepene i hypotesetesting, dvs. hvordan brukes disse verdiene til å konkludere med hva man skal gjøre med nullhypotesen (H_0)? (3 poeng)

- a. P-verdien er et resultat av den statistiske testen vi gjør, og viser oss sannsynligheten for å få det observerte resultatet, eller et mer ekstremt, hvis H_0 er sann.
- b. Signifikansnivået vårt er en forhåndssatt terskelverdi som vi vurderer vår p-verdi opp mot, og er risikoen vi er villige til å ta for å gjøre en type 1-feil.
- c. For å konkludere med hva man skal gjøre med nullhypotesen vurderer man p-verdien opp mot signifikansnivået vi har satt. Om sannsynligheten for å observere dette resultatet ved sann H_0 er lavere enn risikoen vi er villige til å ta for å gjøre en type 1-feil (som regel $p < 0,05$), så forkaster vi nullhypotesen. Er p-verdien høyere enn signifikansnivået, beholder vi nullhypotesen.

Korrelasjon og regresjon

Vi har vært ute og fisket, og fanget mange småfisk. For hver fisk har vi målt vekt i gram (g), samt lengde, høyde og bredde i centimeter (cm). Vi er deretter interessert i å se om det er lengde, høyde eller bredde på fisken som har sterkest sammenheng med fiskens vekt – samt hvor mye disse variablene endres i forhold til hverandre. Signifikansnivået settes til 0,05.

- 1. Hva er forutsetningene for å kunne gjøre en korrelasjonsanalyse? (1 poeng)**
 - a. At det er snakk om to numeriske variabler.
- 2. Vi gjør en korrelasjonsanalyse av alle de aktuelle variablene, og presenterer de i tabellen under. Hvordan korrelerer henholdsvis fiskens lengde, høyde og bredde (cm) med fiskens vekt (g), og hvilken av disse variablene har sterkest korrelasjon med fiskens vekt? (3 poeng)**
 - a. Vi ser at det er sterk til meget sterk positiv sammenheng mellom både fiskens høyde, bredde og lengde med fiskens vekt (korrelasjonskoeffisient fra 0,72 til 0,92). Det vil si at fiskens vekt øker i takt med både fiskens høyde, bredde og lengde. Den sterkeste korrelasjonen ser vi mellom fiskens lengde og vekt, med en korrelasjonskoeffisient på 0,92.

	Vekt (g)	Lengde (cm)	Høyde (cm)	Bredde (cm)
Vekt (g)	1			

Lengde (cm)	0.923044		1	
Høyde (cm)	0.724345	0.703409		1
Bredde (cm)	0.886507	0.878520	0.792881	1

3. Vi har hele tiden vært spesielt interessert i hvordan fiskens lengde i centimeter (X) kan forklare fiskens vekt i gram (Y). Derfor gjør vi en lineær regresjonsanalyse for å se nærmere på dette forholdet, og presenterer resultatene i tabellen under. Hva forteller regresjonskoeffisienten oss her? (2 poeng)

- a. Den bekrefter den positive sammenhengen, og forteller oss i tillegg at for hver økning av fiskens lengde med 1 centimeter så øker fiskens vekt i gjennomsnitt med ca. 28 gram.

Regresjonstatistikk	
Multippel R	0.923043559
R Kvadrat	0.852009412
Justert R Kvadrat	0.851066797
Standardfeil	138.1505079
Observasjoner	159

ANOVA					
	df	SS	MS	F	Significance F
Regresjon	1	17251025.71	17251026	903.8783	<0.001
Residualer	157	2996433.363	19085.56		
Total	158	20247459.07			

	Koeffisienter	Standardfeil	t Stat	P-verdi	Nedre 95%	Øvre 95%
Skjæringspunkt	-490.4005916	31.52561963	-15.5556	<0.001	-552.6696538	-428.1315293
Lengde (cm)	28.4601708	0.946634933	30.06457	<0.001	26.59038774	30.32995387

4. Hva er R Kvadrat i denne modellen, og hva forteller denne verdien oss? (2 poeng)

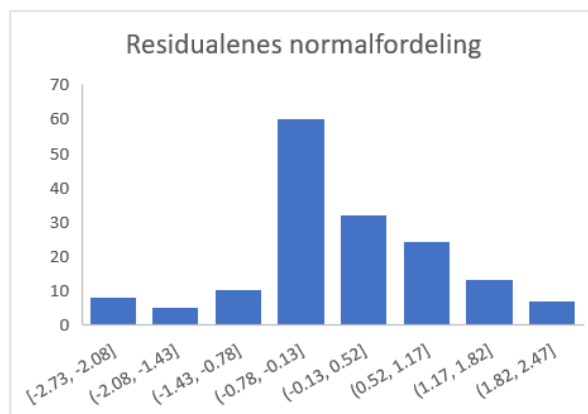
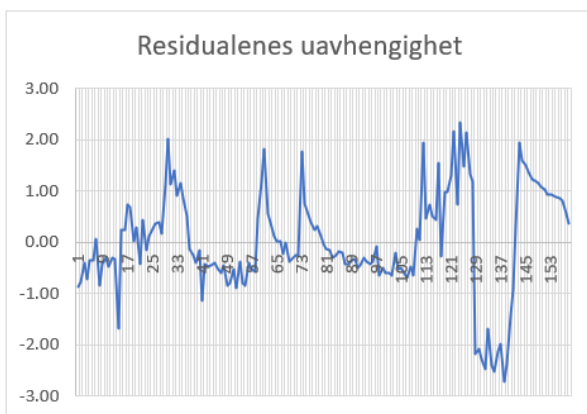
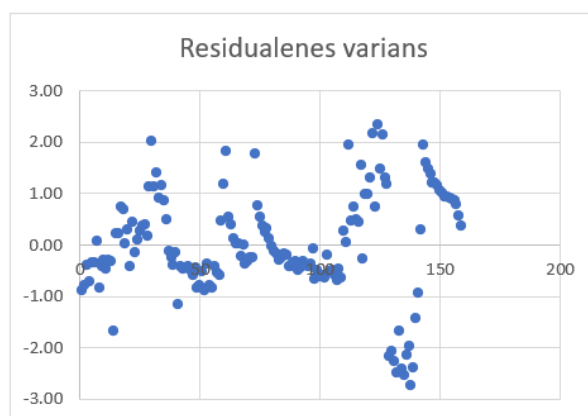
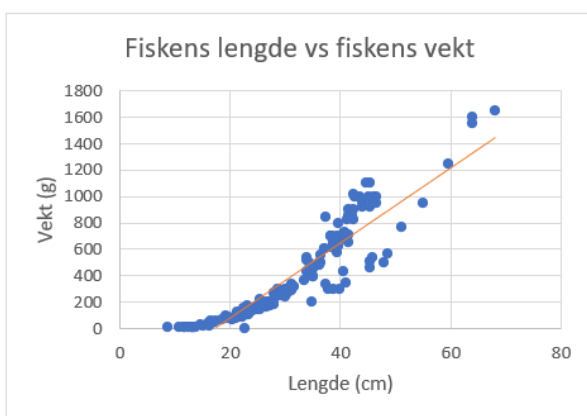
- a. R Kvadrat er her på omtrent 0,85, som betyr at 85 % av variasjonen i fiskens vekt kan forklares med fiskens lengde. Det virker med andre ord som om fiskens lengde er en sterk prediktor av fiskens vekt.

5. Se på figurene under og vurder i hvilken grad forutsetningene for en lineær regresjonsmodell er oppfylt. Hvordan ville du konkludert på sammenhengen mellom fiskens lengde (cm) og fiskens vekt (g) mtp en nullhypotese? (4 poeng)

- a. Vi ser av figurene under at sammenhengen mellom fiskens lengde og fiskens vekt ikke er helt lineær, dvs at en helt rett linje antageligvis ikke er det som passer best til disse dataene. En lineær sammenheng er den viktigste

forutsetningen for å kjøre lineær regresjon. Ellers ser vi at residualenes varians ikke er jevnfordelt og at residualene heller ikke ser helt uavhengige ut fra hverandre. Selv om residualene ser tilnærmet normalfordelt ut, må vi kunne si at forutsetningene for lineær regresjon ikke er oppfylt her.

- b. Som konklusjon ser vi at p-verdien her er godt under signifikansnivået ($p < 0,001$), og vi må derfor forkaste en nullhypotese i dette tilfellet. Det henger imidlertid usikkerhet over disse resultatene da forutsetningene for en lineær regresjon ikke er oppfylt.



Kji-kvadrat (Del 3)

Vi har gjennomført en studie på norske voksne og eldre menn og kvinner, og har blant annet kategorisert alder (40-49 år, 50-64 år, og 65 år og eldre) og utdanningsnivå (Lavt, Middels og Høyt). Deretter er vi interesserte i å se på om det er forskjell i utdanningsnivå på tvers av aldersgruppene i dette utvalget. Signifikansnivået settes til 0,05.

1. Lag nullhypotese og arbeidshypotese for denne problemstillingen. (1 poeng)

- H_0 : Det er ikke en forskjell i utdanningsnivå på tvers av aldersgrupper.
- H_1 : Det er en forskjell i utdanningsnivå på tvers av aldersgrupper.

2. Vi legger våre to variabler i en krysstabell som den under, som viser våre observerte verdier. Tabellen under dette igjen må imidlertid ha forventede verdier for å kunne gjøre Kji-kvadrat i Excel, men denne er tom. Regn ut de forventede verdiene, og rund av til heltall (ingen desimaler). (3 poeng)

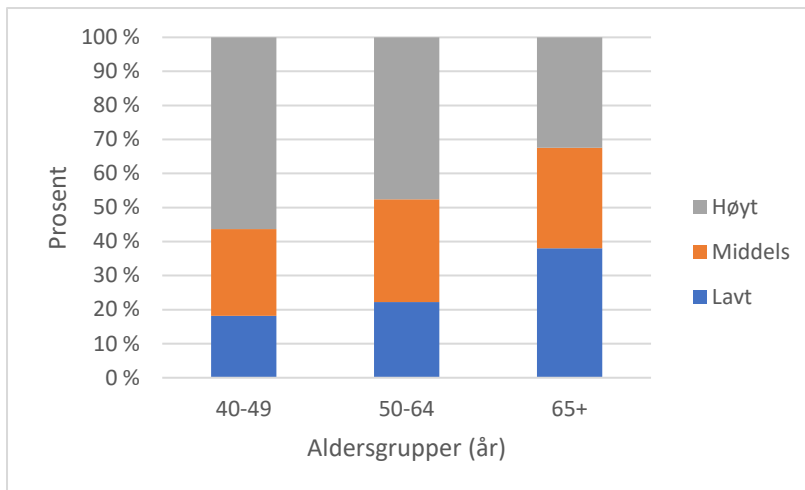
Observerte verdier

Aldersgruppe	Utdanningsnivå			Total
	Lavt	Middels	Høyt	
40-49	10	14	31	55
50-64	14	19	30	63
65+	27	21	23	71
Total	51	54	84	189

Forventede verdier

Aldersgruppe	Utdanningsnivå			Total
	Lavt	Middels	Høyt	
40-49	15	16	24	55
50-64	17	18	28	63
65+	19	20	32	71
Total	51	54	84	189

3. Hva er forutsetningene for å gjøre Kji-kvadrat? Er de tilfredsstilt i dette tilfellet? (1 poeng)
- Forutsetningene for å gjøre Kji-kvadrat (foruten at det er snakk om to kategoriske variabler) er at det forventede antallet i hver celle skal være ≥ 5 . Dette er tilfredsstilt her.
4. Vi lager et stablet stolpediagram for å visualisere utdanningsnivå på tvers av alderskategorier. Regn ut hvor mange prosent i hver alderskategori som har høy utdannelse (rund av til en desimal). (3 poeng)
- Andel med høy utdannelse i alderen 40-49: 56,4%
 - Andel med høy utdannelse i alderen 50-64: 47,6%
 - Andel med høy utdannelse i alderen 65+: 32,4%



5. Etter å ha gjort en Kji-kvadrat test i Excel får vi en p-verdi på 0,044. Konkluder på problemstillingen. (2 poeng)

- a. I vårt utvalg er det større andel med høy utdannelse i de yngre aldersgruppene enn i den eldste aldersgruppen, og p-verdien vår på 0,044 er lavere enn vårt signifikansnivå på 0,05. Med bakgrunn i dette må vi forkaste nullhypotesen i dette tilfellet.