BOKMÅL

Høgskolen i Sørøst-Norge

EKSAMENSFORSIDE
Skriftlig eksamen med tilsyn

Emnekode: 6005	Emnenavn: Statistikk I	
Dato: 02.05.2017	Tid fra / til: 9-13	Ant. timer: 4
Ansv. faglærer: Per Chr. Hagen/Tone K. Grande		
Campus: Bø/Porsgrunn	Fakultet: Handelshøyskolen	
Antall oppgaver: 5	Antall vedlegg:	Ant. sider inkl. forside og vedlegg: 4
Tillatte hjelpemidler (jfr. emnebeskrivelse): Utdelt formelsamling og kalkulator		
Opplysninger om vedlegg:		
Merknader: Alle 10 deloppgaver teller likt ved sensuren.		

Kryss av for type eksamenspapir

Ruter ☐ Linjer ☐

Oppgave 1
En produksjonsbedrift har to maskiner. Vi lar A være utfallet at den ene maskinen virker en hel arbeidsdag, og B er utfallet at den andre maskinen virker en hel arbeidsdag. Vi antar at P(A) = 0.80, P(B) = 0.70 og $P(A \cap B) = 0.60$.

a) Regn ut $P(\overline{A})$, $P(A \cup B)$ og P(B|A). Forklar kort hva disse sannsynlighetene sier oss i den gitte situasjonen. Er A og B uavhengige utfall?

Vi antar at om en eller begge maskinene virker en hel arbeidsdag eller ikke, er uavhengig av hva som skjer på andre arbeidsdager. Hva er sannsynligheten for at minst en av de to maskinene virker hele dagen 3 arbeidsdager på rad?

Oppgave 2
Fra en skoleklasse på $N = 15$ elever skal det trekkes ut tilfeldig 3 forskjellige elever.

a) Hvor mange mulige utvalg fins det ved ikke-ordnet trekning?

I klassen er det 9 gutter og 6 jenter.

Hva er sannsynligheten for at utvalget bare består av gutter og hva er sannsynligheten for at utvalget bare består av jenter?

Vi lar *Y* være antall gutter i utvalget på 3 elever.

b) Vis, ved å regne ut de sannsynlighetene som trengs, at *Y* har følgende sannsynlighetsfordeling (avrundet til 3 desimaler):

x	0	1	2	3
$P(Y=x)$	0.044	0.297	0.475	0.185

Bruk sannsynlighetsfordelingen ovenfor til å regne ut E(Y) og Var(Y). Sammenlign med de verdier som formelsamlingen gir for forventning og varians i denne fordelingen.

Oppgave 3
Lett til moderat mangel på vitamin D er en nokså vanlig tilstand blant personer i den norske befolkningen. I de siste årene har det vært forsket på hvordan lav D-vitaminstatus påvirker risikoen for å utvikle ulike typer helseplager og sykdommer. En D-vitaminkonsentrasjon i blodet lavere enn 50 nanomol pr. liter defineres som lett D-vitaminmangel, mens verdier lavere enn 30 defineres som stor mangel.

Anta at D-vitaminkonsentrasjonen X (nanomol pr. liter) for blodgivere som kommer til sykehuset for å gi blod en bestemt vintermåned er normalfordelt med forventning $\mu = 57$ og standardavvik $\sigma = 22$. Anta dessuten at D-vitaminkonsentrasjonen for forskjellige blodgivere er uavhengige variabler.

a) Finn sannsynligheten for at en tilfeldig valgt blodgiver har en D-vitaminkonsentrasjon lavere enn 50. Regn også ut sannsynligheten $P(50 < X < 100)$. Tegn inn de funne sannsynlighetene som arealer på en skisse av sannsynlighetstettheten til X.

b) Finn sannsynligheten for at den gjennomsnittlige D-vitaminkonsentrasjonen for 25 tilfeldig valgte blodgivere er lavere enn 50. Finn også sannsynligheten (tilnærmet) for at det blant 150 tilfeldig valgte blodgivere denne vintermåneden er flere enn 60 med D-vitaminkonsentrasjon lavere enn 50.

Oppgave 4

Vi betrakter igjen situasjonen i oppgave 3 med den forskjell at forventningen μ, som er gjennomsnittlig D-vitaminkonsentrasjon for blodgivere en bestemt vintermåned, nå er ukjent.

En hypotese går ut på at blodgivere som jevnlig tar kunstig sol (går i solarium) gjennom vintermånedene gjennomsnittlig har høyere D-vitaminkonsentrasjon enn hva man i vintermånedene finner blant blodgivere generelt.

For å undersøke om denne hypotesen kan være riktig, måles D-vitaminkonsentrasjonen hos 20 tilfeldig valgte blodgivere som gjennom vinteren jevnlig tar kunstig sol. Målingene, $X_1, X_2, \ldots, X_{20}$, antas å være normalfordelte med forventning μ og standardavvik σ.

Resultatene er gitt til slutt i oppgaven.

a) Estimer μ og σ.
Finn et 95% konfidensintervall for μ dersom vi antar at $\sigma = 22$.
Hvor mange blodgivere som jevnlig tar kunstig sol, måtte minst ha vært trukket ut for å få et 95 % konfidensintervall for μ med feilmargin mindre enn eller lik 5?

Vi skal teste

$$H_0: \mu = 57 \qquad \text{mot} \qquad H_1: \mu > 57$$

b) Finn den kritiske verdien for testen og angi konklusjonen. Bruk signifikansnivå 5 %. Hva betyr denne konklusjonen i praksis?

c) Gjennomfør testingen i b) når vi antar at σ er ukjent. Hva blir konklusjonen? Bruk også her signifikansnivå 5 %.

Resultater:

X (nanomol pr. liter):	56.7	83.2	58.5	54.7	50.4	95.1	104.3	18.6	77.8
	66.6	66.8	96.7	77.8	77.3	75.6	65.2	22.1	54.6
	68.1	82.9							

$$\bar{X} = 67.65 \qquad \sum(X_i - \bar{X})^2 = 9105.29$$

Oppgave 5

I store byer er det ofte slik at salgsprisen for leiligheter som er sentrumsnære er høyere enn for leiligheter som ligger i noe avstand fra sentrum. For å undersøke sammenhengen mellom salgspris og avstand fra sentrum er det samlet inn data for *salgspris pr. kvadratmeter Y* (i tusen kr) og *avstand fra sentrum x* (i kilometer) for 10 leiligheter som er omsatt i løpet av en måned.

Vi skal studere sammenhengen mellom Y og x ved hjelp av en lineær regresjonsmodell. Vi antar at Y er normalfordelt med forventning $E(Y) = \beta_0 + \beta_1 x$ og kjent standardavvik $\sigma = 5$ tusen kroner. Vi antar dessuten at $Y_1, Y_2, \ldots, Y_{10}$ er uavhengige variabler. Resultatene er gitt til slutt i oppgaven.

a) Estimer β_0 og β_1. Forklar kort med ord hva de estimerte verdiene av β_0 og β_1 uttrykker. Hva er estimert forventet salgspris pr. kvadratmeter for en leilighet som ligger 5 km fra sentrum?

b) Tyder undersøkelsene på at salgsprisen pr. kvadratmeter avtar med mer enn 2 tusen kroner pr kilometer avstand fra sentrum? Formuler dette spørsmålet som en hypotesetest. Finn signifikanssannsyligheten for testen og angi konklusjonen når resultatene er som nedenfor. Bruk signifikansnivå 5 %.

Resultater

x	0	3	4	6	6	8	8	9	10	12
Y	73	57	68	55	50	54	48	52	46	37

$\bar{x} = 6.6 \quad \bar{Y} = 54.0 \quad M = \sum (x_i - \bar{x})^2 = 114.4 \quad \sum (x_i - \bar{x}) Y_i = -303.0$