

Eksamen i STT10011 Statistikk 6/6-2018

Eksamenstid: 4 timer.

Hjelpemidler: Kalkulator av Casio fx-82 serien eller TI BA II plus.

Vedlegg: Formelliste

Alle svar skal begrunnes

Oppgave 1

a) Et fondsselskap tilbyr investeringer i 21 ulike aksjefond. 13 av disse fondene investerer i USA, mens 8 investerer i Europa. En kunde ønsker å investere i 5 ulike aksjefond. Anta at kunden velger tilfeldig blant de 21 aksjefondene.

Hva er sannsynligheten for at kunden ender opp med akkurat 4 europeiske fond?

Hva er sannsynligheten for at kunden ender opp med flere europeiske enn amerikanske fond?

b) En stor vinbutikk i Frankrike selger kun franske og italienske viner. 75% av vinflaskene er franske, mens resten er italienske. Av de franske vinflaskene er 60 % rødvinsflasker og 40 % hvitvinsflasker. Av de italienske vinflaskene er 80 % rødvinsflasker og 20 % hvitvinsflasker.

En vinflaske velges tilfeldig. Hva er sannsynligheten for at dette er en rødvinsflaske?

Det viser seg at flasken inneholder hvitvin. Hva er sannsynligheten for at vinen er fransk?

Oppgave 2

a) Man har kommet fram til følgende sannsynlighetsfordeling for antall biler som en tilfeldig familie eier:

x	0	1	2	3	4	5
$P(X = x)$	0.125	0.495	0.333	0.027	0.019	0.001

Finn forventningen og standardavviket til X .

b) En forsker ønsker å undersøke om det er noen forskjell mellom kvinner og menn når det gjelder det å være venstrehendt.

I en undersøkelse av 450 kvinner og 425 menn fikk man følgende resultater:

	Mann	Kvinne
Venstrehendt	56	42
Høyrehendt	369	408

Tyder disse resultatene på at det er noen forskjell mellom kvinner og menn når det gjelder det å være venstrehendt? Besvar spørsmålet ved å gjennomføre en hypotese-test med 5 % signifikansnivå.

Oppgave 3

a) En psykolog påstår at 15% av befolkningen lider av flyskrekk. Gå ut fra at denne påstanden er sann. En bedrift har 30 ansatte. Hva er sannsynligheten for at akkurat 7 av de ansatte har flyskrekk?

b) La p være sannsynligheten for at en tilfeldig valgt person har flyskrekk. Vi skal nå gå ut fra at p er ukjent. I en undersøkelse av 400 tilfeldig valgte personer, viste det seg at 66 har flyskrekk. Lag et 95 % -konfidensintervall for p .

c) La situasjonen være som i b). Bruk resultatet av undersøkelsen nevnt i b) til å teste om det er grunnlag til å tro at p er større enn 0.15. Bruk 5 % signifikansnivå. Finn P -verdien til testen.

Hva ville konklusjonen blitt hvis en hadde valgt 1 % signifikansnivå? Begrunn svaret.

d) Finn styrken i alternativet $p = 0.20$ ved testen beskrevet i b). Forklar hva styrken betyr i dette tilfellet.

Oppgave 4

a) En bedrift har en dagsproduksjon X som er normalfordelt med forventning

$\mu = 25.6$ tonn og standardavvik $\sigma = 5.1$ tonn.

Finn sannsynligheten for at det en dag produseres mindre enn 18 tonn av varen.

Finn også sannsynligheten for at dagsproduksjonen en tilfeldig dag ligger mellom 24 og 30 tonn.

b) Man prøver ut en ny produksjonsmetode. Dagsproduksjonen med den nye produksjonsmetoden antas å være normalfordelt med ukjent forventning μ_1 og ukjent standardavvik σ_1 .

Man registrerer dagsproduksjonen på 20 dager med den nye produksjonsmetoden, og man får da en gjennomsnittlig dagsproduksjon $\bar{X} = 27.8$ tonn og et empirisk standardavvik $S_X = 6.0$ tonn.

Gjennomfør en hypotesetest for påstanden om at μ_1 er større enn 25.6 tonn. Bruk 5 % signifikansnivå.

Lag også et 95 % - konfidensintervall for μ_1 .

c) Anta nå at σ_1 er kjent og lik standardavviket i a), dvs. $\sigma_1 = 5.1$. Hvor mange observasjoner må man da ha for å få et 95 % - konfidensintervall for μ_1 med lengde høyst lik 2?

Oppgave 5

Eieren av en kiosk ønsker å analysere sammenhengen mellom solgt kvantum av en bestemt type pølse og prisen på pølsene. Tabellen viser 8 tilfeldige observasjonspaar av $X =$ prisen på pølse og $Y =$ antall solgte pølser per dag.

X	40	46	52	58	64	70	76	82
Y	440	420	380	370	380	340	300	260

I beregningene nedenfor kan du velge å bruke noen av følgende mellomregninger:

$$\bar{X} = 61, \bar{Y} = 361.25, S_X = 14.6969, S_Y = 59.6268$$

$$\sum_{i=1}^8 (X_i - \bar{X})^2 = 1512 \qquad \sum_{i=1}^8 (Y_i - \bar{Y})^2 = 24887.5$$

$$\sum_{i=1}^8 (X_i - \bar{X}) \cdot (Y_i - \bar{Y}) = -5910 \quad \sum_{i=1}^8 (Y_i - \hat{Y}_i)^2 = 1788.42$$

a) Finn korrelasjonskoeffisienten r .

Hva forteller r oss om forholdet mellom X og Y ?

Finn også den estimerte regresjonslinjen $\hat{Y} = \hat{\alpha} + \hat{\beta} \cdot X$.

Anslå antall solgte pølser når prisen på pølser er 60.

b) Vi antar at kravene til en standard regresjonsmodell er oppfylt. Test hypotesen $\beta \neq 0$, der β er stigningstallet til linjen.

Bruk 5 % signifikansnivå.

Lag også et 95 % - konfidensintervall for β .

Forklar sammenhengen mellom hypotesetesten og konfidensintervallet.