

Institutt for matematiske fag

Eksamensoppgåve i **ST0103 Brukarkurs i statistikk**

Fagleg kontakt under eksamen: Øyvind Bakke

Tlf: 73 59 81 26, 990 41 673

Eksamensdato: august 2015

Eksamenstid (frå–til): 9.00–13.00

Hjelpemiddelkode/Tillatne hjelpemiddel: Gult A4-ark med egne handskrivne notat, bestemd enkel kalkulator, *Tabeller og formler i statistikk* (Tapir forlag), *Matematisk formelsamling* (K. Rottmann)

Annan informasjon:

I vurderinga tel kvart av dei ti bokstavpunkta likt.

Alle svara skal grunngjevast (t.d. ved at mellomrekning blir tatt med eller ved tilvising til teori eller døme frå pensum).

Målform/språk: nynorsk

Sidetal: 2

Sidetal vedlegg: 0

Kontrollert av:

Dato

Sign

Oppgåve 1

Talet på egg, X , i eit tilfeldig valt reir av ein fugleart, har denne sannsynsfordelinga:

$$P(X = x) \begin{array}{l} \left| \begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \\ 0,09 & 0,38 & 0,25 & 0,20 & 0,06 & 0,02 \end{array} \right. \end{array}$$

- a) Finn forventningsverdien og standardavviket til X .
- b) Finn sannsynet for at eit tilfeldig valt reir av denne arten inneheld 4 eller fleire egg. Finn det vilkårsbundne (betinga) sannsynet for at eit tilfeldig valt reir av denne arten inneheld 6 egg gitt at det inneheld 4 eller fleire egg.

Vi undersøker 50 tilfeldig valde reir av fuglearten. La Y vere talet av desse som har 4 eller fleire egg.

- c) Kva er forventningsverdi og standardavvik til Y ? Finn eit tilnærma sannsyn for at 10 eller fleire av dei 50 reira minst har 4 egg.

Det finst (nøyaktig) to underartar av denne arten, A og B . 5% av reira av underart A inneheld berre eitt egg, medan 15% av reira av underart B berre inneheld eitt egg.

- d) Kor stor del av reira høyrer til underart A ?

Oppgåve 2

I eit laboratorium blir eit løysingsmiddel brukt som skal ha $pH \leq 7,45$. Det blir tatt n prøvar av eit parti av løysingsmiddelet, og i kvar prøve blir pH målt. Målingane, X_1, X_2, \dots, X_n , er uavhengige og normalfordelte med forventningsverdi μ , som er pH i løysingsmiddelet, og standardavvik 0,05 (som botnar i måleutryggleik).

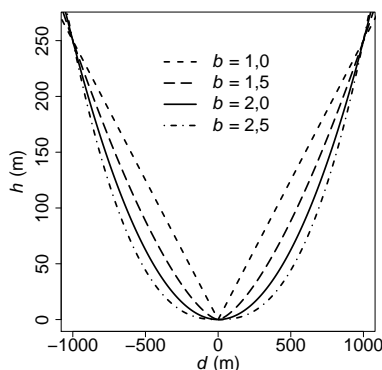
Ein utfører ein hypotesetest med nullhypotese $\mu \leq 7,45$ mot den alternative hypotesen $\mu > 7,45$. Dersom nullhypotesen blir forkasta med signifikansnivå 0,05, blir partiet kassert.

- a) Foreslå korleis ein slik hypotesetest basert på gjennomsnittet \bar{X} av målingane kan gjennomførast. Kva blir konklusjonen dersom det blir tatt $n = 20$ prøvar og gjennomsnittet av målingane er 7,47?
- b) Kva er sannsynet for at nullhypotesen blir forkasta dersom det blir tatt $n = 20$ prøvar og pH i løysingsmiddelet er $\mu = 7,47$?
- c) Anta at pH i løysingsmiddelet er $\mu = 7,47$. Kor mange prøvar, n , må ein ta for at sannsynet for at nullhypotesen blir forkasta skal vere større enn 0,8?

Oppgåve 3

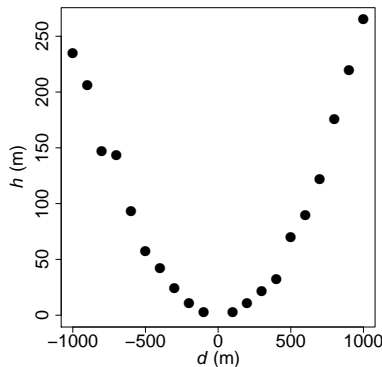
U-dalar er skura ut av brear frå istida, og har eit bogent tverrsnitt, i motsetnad til V-dalar, som er gravne ut av vatn og har eit V-forma tverrsnitt. Vi antar at tverrsnittet av ein dal kan skildrast ved likninga $h = a|d|^b$ i eit kartesisk (d, h) -koordinatsystem med vassrett d -akse (på tvers av dalen), loddrett h -akse og origo i dalbotnen, der a og b er ukjende parametrar. Då vil ein V-dal ha b nær 1 (høgda over dalbotnen er proporsjonal med avstanden frå dalbotnen), medan ein U-dal vil ha ein større b , til dømes nær 2 (som gir eit parabelforma tverrsnitt).

Til høgre er det vist tverrsnitt for fire ulike verdiar av b (a er ulik for dei fire tverrsnitta).



Frå eit tverrsnitt av ein dal har vi 20 målingar (d_i, h_i) , $i = 1, 2, \dots, 20$ (sjå plott til høgre), og vi ønsker å estimere a og b slik at punkta (d_i, h_i) ligg i nærleiken av grafen til $h = a|d|^b$.

- a) Vis at vi ved å transformere (omforme) d til $x = \ln|d|$ og h til $y = \ln h$ får ein lineær samanheng mellom x og y , $y = \alpha + \beta x$, slik at vi kan estimere parametrane ved lineær regresjon. Vis at $b = \beta$, og finn samanhengen mellom a og α .



Vi måler d og h i meter. For dei transformerte dataa har vi $\bar{x} = 6,12$, $\bar{y} = 3,96$, $\sum_{i=1}^{20} (x_i - \bar{x})(y_i - \bar{y}) = 19,25$ og $\sum_{i=1}^{20} (x_i - \bar{x})^2 = 9,67$.

- b) Rekn ut estimata av α og β gitt ved minste kvadrats metode.
- c) Finn eit 95%-konfidensintervall for β . Kva føresetnader ligg til grunn for regresjonsmodellen for at dette intervallet skal vere rett? Bør dalen klassifiserast som U-dal eller V-dal ut frå konfidensintervallet? Det blir oppgitt at *error sum of squares* er $SS_E = 0,156$.