

Institutt for matematiske fag

## Eksamensoppgave i **ST0103 Brukerkurs i statistikk**

**Faglig kontakt under eksamen:** Øyvind Bakke

**Tlf:** 73 59 81 26, 990 41 673

**Eksamensdato:** august 2015

**Eksamenstid (fra–til):** 9.00–13.00

**Hjelpemiddelkode/Tillatte hjelpemidler:** Gult A4-ark med egne håndskrevne notater, bestemt enkel kalkulator, *Tabeller og formler i statistikk* (Tapir forlag), *Matematisk formelsamling* (K. Rottmann)

**Annen informasjon:**

I vurderingen teller hvert av de ti bokstavpunktene likt.

Alle svar skal begrunnes (f.eks. ved at mellomregning tas med eller ved henvisning til teori eller eksempler fra pensum).

**Målform/språk:** bokmål

**Antall sider:** 2

**Antall sider vedlegg:** 0

**Kontrollert av:**

---

Dato

Sign



**Oppgave 1**

Antall egg,  $X$ , i et tilfeldig valgt reir av en fugleart, har denne sannsynlighetsfordelingen:

$x$	1	2	3	4	5	6
$P(X = x)$	0,09	0,38	0,25	0,20	0,06	0,02

- Finns forventningsverdien og standardavviket til  $X$ .
- Finns sannsynligheten for at et tilfeldig valgt reir av denne arten inneholder 4 eller flere egg. Finn den betingede sannsynligheten for at et tilfeldig valgt reir av denne arten inneholder 6 egg gitt at det inneholder 4 eller flere egg.

Vi undersøker 50 tilfeldig valgte reir av fuglearten. La  $Y$  være antallet av disse som har 4 eller flere egg.

- Hva er forventningsverdi og standardavvik til  $Y$ ? Finn en tilnærmet sannsynlighet for at 10 eller flere av de 50 reirene minst har 4 egg.

Det fins (nøyaktig) to underarter av denne arten,  $A$  og  $B$ . 5% av reirene av underart  $A$  inneholder bare ett egg, mens 15% av reirene av underart  $B$  bare inneholder ett egg.

- Hvor stor andel av reirene tilhører underart  $A$ ?

**Oppgave 2**

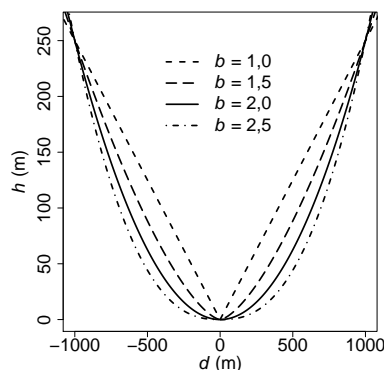
I et laboratorium brukes et løsningsmiddel som skal ha  $pH \leq 7,45$ . Det tas  $n$  prøver av et parti av løsningsmiddelet, og i hver prøve måles  $pH$ . Målingene,  $X_1, X_2, \dots, X_n$ , er uavhengige og normalfordelte med forventningsverdi  $\mu$ , som er  $pH$  i løsningsmiddelet, og standardavvik 0,05 (som skyldes måleusikkerhet).

En utfører en hypotesetest med nullhypotese  $\mu \leq 7,45$  mot den alternative hypotesen  $\mu > 7,45$ . Hvis nullhypotesen forkastes med signifikansnivå 0,05, kasseres partiet.

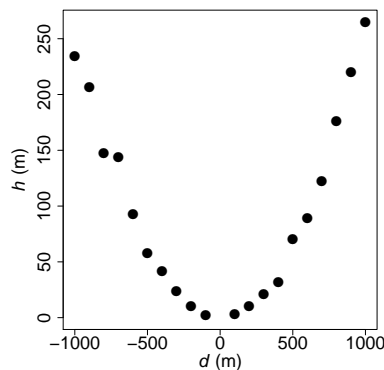
- Foreslå hvordan en slik hypotesetest basert på gjennomsnittet  $\bar{X}$  av målingene kan gjennomføres. Hva blir konklusjonen hvis det tas  $n = 20$  prøver og gjennomsnittet av målingene er 7,47?
- Hva er sannsynligheten for at nullhypotesen blir forkastet hvis det tas  $n = 20$  prøver og  $pH$  i løsningsmiddelet er  $\mu = 7,47$ ?
- Anta at  $pH$  i løsningsmiddelet er  $\mu = 7,47$ . Hvor mange prøver,  $n$ , må tas for at sannsynligheten for at nullhypotesen forkastes skal være større enn 0,8?

### Oppgave 3

U-daler er skurt ut av breer fra istida, og har et buet tverrsnitt, i motsetning til V-daler, som er gravd ut av vann og har et V-formet tverrsnitt. Vi antar at tverrsnittet av en dal kan beskrives ved likningen  $h = a|d|^b$  i et kartesisk  $(d, h)$ -koordinatsystem med vannrett  $d$ -akse (på tvers av dalen), loddrett  $h$ -akse og origo i dalbunnen, der  $a$  og  $b$  er ukjente parametre. En V-dal vil da ha  $b$  nær 1 (høyde over dalbunnen er proporsjonal med avstand fra dalbunnen), mens en U-dal vil ha en større  $b$ , for eksempel nær 2 (som gir et parabelformet tverrsnitt). Til høyre er det vist tverrsnitt for fire forskjellige verdier av  $b$  ( $a$  er forskjellig for de fire tverrsnittene).



Fra et tverrsnitt av en dal har vi 20 målinger  $(d_i, h_i)$ ,  $i = 1, 2, \dots, 20$  (se plott til høyre), og vi ønsker å estimere  $a$  og  $b$  slik at punktene  $(d_i, h_i)$  ligger i nærheten 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 en lineær sammenheng mellom  $x$  og  $y$ ,  $y = \alpha + \beta x$ , slik at vi kan estimere parametrene ved lineær regresjon. Vis at  $b = \beta$ , og finn sammenhengen mellom  $a$  og  $\alpha$ .

Vi måler  $d$  og  $h$  i meter. For de transformerte dataene 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) Regn ut estimatene av  $\alpha$  og  $\beta$  gitt ved minste kvadraters metode.
- c) Finn et 95 %-konfidensintervall for  $\beta$ . Hvilke forutsetninger ligger til grunn for regresjonsmodellen for at dette intervallet skal være riktig? Bør dalen klassifiseres som U-dal eller V-dal ut fra konfidensintervallet? Det oppgis at *error sum of squares* er  $SS_E = 0,156$ .