**Noregs teknisk–naturvitskaplege universitet**
**Institutt for matematiske fag**



Nynorsk

Fagleg kontakt under eksamen:

# EKSAMEN I EMNE TMA4245 STATISTIKK

7. august 2008
Tid: 09:00–13:00

Hjelpemiddel: *Tabeller og formler i statistikk*, Tapir Forlag
K. Rottmann: *Matematisk formelsamling*
Kalkulator HP30S
Gult, stempla A5-ark med eigne handskrevne notat.

## Oppgåve 1

Vi skal i denne oppgåva sjå på høgdefordelinga til menn og kvinner. Anta at høgda til menn er normalfordelt med forventningsverdi $\mu_M = 179$ og varians $\sigma_M^2 = 6^2$, og at høgda til kvinner er normalfordelt med forventningsverdi $\mu_K$ og varians $\sigma_K^2$

**a)** Finn sannsynet for at ein tilfeldig valt mann er over 185 cm.

Finn sannsynet for at ein mann er over 185 cm gitt at han er over 179 cm.

**b)** La $\mu_K = \beta \cdot \mu_M$, der $\beta$ er ein ukjent parameter vi ønsker å estimere. Anta at vi har høgdedata fråeit tilfeldig utval på $n = 5$ kvinner; $X_i \sim N(\mu_K, \sigma_K^2)$, $i = 1, 2, \ldots, n$.

Som estimator for $\beta$ veljer vi $\hat{\beta} = \dfrac{\sum_{i=1}^n X_i}{n\mu_M}$

Er $\hat{\beta}$ forventningsrett?

Utlei eit uttrykk for eit 95% konfidensintervall for $\beta$, og finn talsvar når vi frå data får $\bar{x} = \sum_{i=1}^n x_i/n = 167.5$ og $s^2 = \dfrac{1}{n-1}\sum_{i=1}^n (x_i - \bar{x})^2 = 5.1^2$.

**Oppgåve 2**

Ei næringsmiddelverksemd har eit laboratorium for testing av kvaliteten av produkta sine. For testing av mengde pr. volumeining av eit bestemt sporstoff i eit av produkta, bruker verksenda to måleapparat, som vi kan kalle A og B. A er det mest nøyaktige av dei to, og det er og det raskaste. Ulempa med A er at det er dyrt i bruk. Versemda har derfor etablert ei testprosedyre som går ut på å ta ut to tilfeldige utval frå ein gitt produksjonsserie, der vi kan anta at mengde pr. volumeining, kalla $\mu$, er konstant. Det eine utvalet, som er av storleik $n$, blir testa i A. Etter denne testinga i A har ein då eit tilfeldig utvalg $x_1, \ldots, x_n$ med måleresultater. Det andre utvalet, som er av storleik $2n$, blir testa i B. Det resulterer i eit uavhengig tilfeldig utval $y_1, \ldots, y_{2n}$ av storleik $2n$ med måleresultater etter testing i B.

Det tilfeldige utvalget $x_1, \ldots, x_n$ kan ein sjå på som eit utfall av $n$ stokastiske variablar $X_1, \ldots, X_n$, som er uavhengige og normalfordelte med forventningsverdi $\mu$ og varians $\sigma^2$, der $\sigma^2$ reflekterer A sin nøyaktigheit. Tilsvarande, $y_1, \ldots, y_{2n}$ kan ein sjå på som utfall av $2n$ stokastiske variablar $Y_1, \ldots, Y_{2n}$, som er uavhengige og normalfordelte med forventningsverdi $\mu$ og varians $4\sigma^2$, og dei er uavhengige av $X_i$, $i = 1, \ldots, n$.

**a)** Mengde pr. volumeining $\mu$ skal estimerast på grunnlag av dei tilsamen $3n$ målingane som er gjort, og ein mogleg estimator er

$$\hat{M} = \frac{1}{2}\left(\overline{X} + \overline{Y}\right) \qquad \left(\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i\,, \quad \overline{Y} = \frac{1}{2n}\sum_{j=1}^{2n} Y_j\,\right)$$

Kva er forventningsverdi og varians til denne estimatoren?

**b)** Vis at sannsynsmaksimerings-estimatoren (SME) for $\mu$ er gitt som

$$M^* = \frac{1}{3}\left(2\overline{X} + \overline{Y}\right)\ ,$$

og finn forventningsverdi og varians.

Kven av dei to estimatorane $\hat{M}$ og $M^*$ ville du tilrå? Grunngje svaret.

**c)** Anta i dette punktet at $\sigma^2$ er kjent, og lik 1.0. Ein skal no teste hypotesa

$$H_0\ :\ \mu = \mu_0 = 100\,,$$

mot

$$H_1\ :\ \mu < \mu_0\,.$$

på signifikansnivå $\alpha = 0.05$. Foreslå ein testobservator, og angje testen sitt kritiske område.

Kva blir konklusjonen når $n = 4$ og måleresultatene er:

$X_i$: 100.3, 100.8, 99.5, 98.8; som gir $\overline{X} = 100.4$

$Y_j$: 100.1, 98.7, 99.2, 102.3, 103.0, 97.5, 96.6, 103.1; som gir $\overline{Y} = 100.5$

**d)** Anta i dette punktet at $\sigma^2$ er ukjent. Følgjende to estimatorar blir innført,

$$S_1^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2 \qquad \left(\mathrm{E}[S_1^2] = \sigma^2\right)$$

og

$$S_2^2 = \frac{1}{2n-1} \sum_{j=1}^{2n} \left(Y_j - \overline{Y}\right)^2 \qquad \left(\mathrm{E}[S_2^2] = 4\sigma^2\right)$$

Desse to blir kombinert i ein ny stokastisk variabel,

$$V = \frac{(n-1)S_1^2}{\sigma^2} + \frac{(2n-1)S_2^2}{4\sigma^2}\,.$$

Grunngje kvifor $V$ blir $\chi^2$-fordelt. Kvor mange fridomsgrader har $V$?

Foreslå ein testobservator for $\mu$ basert på $M^*$ og $V$ og angje fordeling.

## Oppgåve 3

I semifinala mellom Frankrike og Tyskland i eit verdsmesterskap i fotball er resultatet uavgjort etter ekstraomgangar.

Vi definerer ein runde i straffesparkkonkuransen til å vere at kvart lag skyt en straffe kvar.

I første del av straffesparkkonkurransen er det 5 rundar, 5 straffespark frå kvart lag, og laget som skårer flest gongar vinn.

Dersom laga skårer like mange gongar i første del, går ein over til andre del av straffesparkkonkurransen. No blir det spelt ein og ein runde: Kvart lag har ein straffe kvar. Dersom det eine laget skårer og det andre bommar, har vi en vinner. Ellers blir det spilt ein ny runde (kvart lag får ein straffe) inntil vi har ein vinnar.

Anta at dei tyske spelarane har sannsyn $p_T = 0.80$ for å skåre på straffe, at dei franske spelarane har sannsyn $p_F = 0.70$ for å skåre, og at utfalla av straffesparka er uavhengige av kvarandre.

**a)** Hva er sannsynet for at stillinga blir 5-5 etter første del?

Hva er sannsynet for at stillinga blir 3-3 etter første del?

Sett opp uttrykk for / algoritme for korleis ein kan finne sannsynet $p_{D1-lik}$ for at laga står likt etter første del?

Du treng ikkje finne talsvar, og kan seinare i oppgåva bruke at $p_{D1-lik} = 0.27$

**b)** Anta i dette punktet at vi er komme til del 2 av straffesparkkonkuransen.

Hva er sannsynet $p_{V1|D2}$ for at vi har ein vinnar etter første runde i del to?

Gitt at vi har ein vinner etter første runde i del 2, kva er sannsynet $p_{T|V1,D2}$ for at dette er Tyskland?

Kva er sannsynet $p_{T|D2}$ for at Tyskland vinn konkurransen (gitt at vi er komme til del 2)?

**c)** La $X$ vere antall runder i del 2 til og med runden det blir kåra ein vinnar i.

Kva fordeling har $X$? Grunngje svaret og oppgje parameter/re.

Kva er forventningsverdien og variansen i antall rundar (i del 2) til det blir kåra ein vinnar?

**d)** Vi antar at det alltid blir spelt 5 rundar i del 1. Kva er forventa totalt antall rundar i straffesparkkonkurransen (del 1 og evt. del 2) til det blir kåra ein vinnar?

Finn og variansen av det totale antalet rundar i straffesparkkonkurransen (del 1 og evt. del 2) inntil det blir kåra ein vinnar.