

ECON2130: EKSAMEN 2010 VÅR - UTSATT PRØVE

Oppgave 1

Anta X er poissonfordelt med parameter $\lambda = 5$ (skrevet kort, $X \sim \text{pois}(5)$, jfr. definisjon 5.8 i Løvås med $t = 1$).

A. (i) Finn $P(X = 5)$ og $P(X \geq 5)$, for eksempel basert på tabell D.2 i Løvås.

(ii) Anta vi vet om en observasjon av X at den ikke er større enn 5. Hva er da sannsynligheten for at den er lik 5? (**Hint:** Finn $P(X = 5 | X \leq 5)$).

B. Beregn $P(X = 5)$ tilnærmet ved bruk av normaltilnærmelsen til poisson-fordelingen med heltallskorreksjon.

C. Anta at 10 stokastiske variable, X_1, X_2, \dots, X_{10} , er uavhengige og identisk poisson-fordelte, alle med parameter $\lambda = 5$. Beregn

(i) sannsynligheten for at ingen av disse 10 X_i -ene får verdien 5,

(ii) sannsynligheten for at akkurat 2 av de 10 X_i -ene blir 5,

(iii) forventet antall av de 10 X_i -ene som blir 5.

Oppgave 2

(Som oppgave 2.19 i Løvås - reprodusert og modifisert her:)

La x_1, x_2, \dots, x_n være n observasjoner av en variabel. Den vanlige formelen for (empirisk)

variens som er gitt i definisjon 2.3 i Løvås, er $S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$ der $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$. Denne

formelen kan være tungvint å bruke hvis du skal regne for hånd. Vis at variansen, S^2 , også kan beregnes med formelen

$$S^2 = \frac{1}{n-1} \left(\sum x_i^2 - n\bar{x}^2 \right)$$

Hint: Start med definisjonen (jmfør def. 2.3 i Løvås), og bruk regneregler for sum (for eksempel regel 2.2 i Løvås). Tenk på \bar{x} som en konstant du kan kalle b .

Oppgave 3

Det er mye tilfeldigheter som spiller inn i fotballkamper mellom likeverdige lag. Det kan for eksempel virke som at antall mål som et lag klarer å skåre i en slik kamp er poisson-fordelt. Vi skal i denne oppgaven se litt nærmere på denne muligheten ved å studere resultatene fra gruppespillet under VM i fotball i Frankrike i 1998 (der også Norge var med og slo Brasil 2 - 0).

I gruppespillet var det 32 lag som alle spilte 3 kamper - altså i alt 96 forsøk der "forsøk" viser til det at et lag deltar i en kamp. Hver kamp omfatter således to forsøk, et for hvert lag. I hvert forsøk registreres antall mål som det aktuelle laget skårer. Tabell 1 viser resultatene for de 96 forsøkene. For eksempel, 26 forsøk resulterte i 0 mål. (Et av disse forsøkene var Brasils kamp mot Norge som ga 0 mål. Norge i samme kamp representerte et annet forsøk som ga 2 mål.)

Tabell 1

Antall mål	0	1	2	3	4	5	6	Sum
Frekvens	26	34	24	8	1	2	1	96

- A.** La x_1, x_2, \dots, x_n , der $n = 96$, være observasjonene fra de 96 forsøkene oppsummert i tabell 1. Forklar hvordan du ville regne for å komme fram til

$$\sum_{i=1}^{96} x_i = 126 \quad \text{og} \quad \sum_{i=1}^{96} x_i^2 = 304$$

- B.** La X_i betegne antall mål skåret i i -te forsøk, $i = 1, 2, \dots, n = 96$, tolket som en stokastisk variabel. Som modell antar vi at X_1, X_2, \dots, X_n , er uavhengige og identisk poisson-fordelte, $X_i \sim \text{pois}(\lambda)$, der $\lambda = E(X_i)$ er en ukjent parameter.

(i) Vis at

$$\hat{\lambda} = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

er en forventningsrett estimator for λ .

(ii) Finn variansen til $\hat{\lambda}$ uttrykt ved λ .

(iii) Beregn estimatet basert på $\hat{\lambda}$ og data i tabell 1.

C. Det kan vises (som du ikke trenger å gjøre) at $\frac{\hat{\lambda} - \lambda}{\sqrt{\hat{\lambda}}} \sqrt{n}$ er tilnærmet standard normalfordelt, $N(0, 1)$, når n er stor ($n = 96$ burde være tilstrekkelig). Bruk dette til å utlede et (tilnærmet) 95% konfidensintervall for λ . Beregn intervallet ut fra data i tabell 1.

D. Tatt i betraktning at ferdighetene til de forskjellige lagene tross alt er forskjellige, kan forutsetningen i punkt B om uavhengige og identisk poisson-fordelte variable synes tvilsom. Vi skal derfor se på om data kan gi noen indikasjoner som taler imot denne antakelsen.

Anta $X \sim \text{pois}(1,3)$, der parameterverdien $\lambda = 1,3$ ligger nær den estimerte verdien i punkt B. Beregn sannsynlighetene $P(X = x)$ for $x = 0, 1, 2, 3$, samt $P(X \geq 4)$.

Sammenlign disse 5 sannsynlighetene med tilsvarende relative frekvenser beregnet ut fra tabell 1 og kommenter.

E. **Innledning.** Vi kan også sette opp en formell statistisk test for antakelsen i punkt B. La null-hypotesen, H_0 , være at X_1, X_2, \dots, X_n er uavhengige og identisk poisson-fordelte med parameter λ . Hvis H_0 er riktig, vil X_i -ene ha samme forventning og varians, λ . Nå er $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$, generelt en forventningsrett estimator for $\text{var}(X_i) = \sigma^2$. Siden $\sigma^2 = \lambda$ under H_0 , vil både $\hat{\lambda} = \bar{X}$ og S^2 estimere samme parameter hvis H_0 er sann. Dette impliserer at, hvis H_0 er sann, så vil variabelen $V = \frac{S^2}{\bar{X}}$ variere i nærheten av 1. Vi kan derfor benytte V som testobservator og forkaste H_0 hvis V får en verdi tilstrekkelig langt vekk fra 1. Med andre ord, forkast H_0 hvis $V \leq c_1$ eller $V \geq c_2$, der c_1, c_2 er passende kritiske verdier.

Oppgave.

(i) Bestem de kritiske verdiene c_1, c_2 slik at testen får signifikansnivå (tilnærmet) 5%.

[Hint. Det kan vises (som du ikke trenger å gjøre), at, hvis H_0 er sann, så er V tilnærmet normalfordelt med forventning 1 og varians $\frac{1+1/\lambda}{n}$, når n er stor,

der den ukjente λ i variansen kan erstattes med estimatet, $\hat{\lambda}$, eller 1,3 om du vil. Bruk dette til å bestemme de kritiske verdiene.]

- (ii) Gjennomfør testen ut fra data i tabell 1 og formuler en konklusjon.