

TENTAMEN: Matematisk statistik och diskret matematik (MVE050/MSG810).

Statistik för fysiker (MSG820)

Tid och plats: Fredagen den 18 december 2009, kl. 08.30–12.30, Väg och vatten-salar.

Jour: Oscar Hammar, tel 0708-300715.

Tillåtna hjälpmedel: Chalmersgodkänd räknare och Beta.

Betygsgränser: Chalmers: 3: 12 poäng, 4: 18 poäng, 5: 24 poäng. GU: G: 12 poäng VG: 21 poäng. Maximalt antal poäng är 30.

1. (3p) Låt $P(A) = 0.5$ och $P(B) = 0.4$. Beräkna $P(A|B)$ då
 - a) A och B är oberoende
 - b) $P(B|A) = 0.35$
 - c) $P(A \cup B) = 0.7$
2. (3p) Låt X vara en stokastisk variabel och $\mu_X = \mathbf{E}[X]$. Låt (X_1, \dots, X_{10}) vara ett stickprov från fördelningen för X . En skattare är en funktion av ett stickprov. Betrakta följande två skattare av μ_X :

- $W_1(X_1, \dots, X_{10}) = \frac{1}{10} \sum_{i=1}^{10} X_i$
- $W_2(X_1, \dots, X_{10}) = X_1$

Vi vet att W_1 är en väntevärdesriktig skattare av μ_X .

- a) Är W_2 en väntevärdesriktig skattare av μ_X ? Motivera ditt svar.
 - b) Vilken av W_1 och W_2 är att föredra? Motivera ditt svar.
3. (4p) Rolf vill undersöka hur stor andel av en population av möjliga datorköpare som tänker köpa ny dator inom det närmaste halvåret. Han tar ett stickprov av storlek 100 av populationen. Av dessa uppger 6 att de planerar att köpa en ny dator inom det närmast halvåret.
 - a) Beräkna ett konfidensintervall åt Rolf med valfri konfidensgrad för andelen i populationen som tänker köpa en ny dator inom det närmaste halvåret baserat på den insamlade datan
 - b) Kan du garantera konfidensgraden för intervallet du givit i deluppgift a? Motivera ditt svar.
 4. (4p) I snitt inkommer 7 mail/dygn till Kurts mailbox. Kurt har många internationella vänner så intensiteten för inkommande mail är ungefär konstant fördelad över dygnet. Kurt antar att inkommande mail följer en poissonprocess.

Kurt kollar sin mail det sista han gör innan han går och lägger sig och det första han gör när han vaknar. Han sover 8h/natt.

 - a) Vad är enligt Kurts modell sannolikheten att han inte får ett enda mail under natten?
 - b) Vad säger Kurts modell om sannolikheten att få noll mail minst två av veckans sju nätter?

5. (4p) En forskargrupp vill undersöka vattentemperaturökningen i en å 30 meter nedströms från ett kärnkraftverk. Vid slumpvisa tillfällen mäts skillnaden i vattentemperatur mellan platsen 30 meter nedströms och på en plats uppströms från kraftverket där vattnet antas vara opåverkat av kraftverket.

Enligt gällande bestämmelser får medeltemperaturökningen ej överstiga $0.5^\circ C$. Forskarnas 6 observationer av vattentemperaturökningen presenteras nedan:

0.9 1.1 1.1 0.4 1.0 1.3

- a) Inför lämpliga beteckningar och gör lämpliga fördelningsantaganden.
b) Testa nollhypotesen att medeltemperaturökningen inte överstiger $0.5^\circ C$ mot mothypotesen att medeltemperaturökningen överstiger $0.5^\circ C$.
c) Hur kan man öka ett tests styrka utan att påverka signifikansnivån?
6. (4p) Låt X vara en stokastisk variabel med $\mathbf{E}[X] = \mu_X$. Härled ett 95%-igt konfidensintervall för μ_X baserat på ett stickprov (X_1, \dots, X_n) från fördelningen för X under antagandet att $X \sim N(\mu_X, \sigma_X)$, där σ_X är känd.
- Hjälp: Vilken punktskattare används för att skatta μ_X ? Om $X \sim N(\mu_X, \sigma_X)$, vad har då punktskattaren för μ_X för fördelning.

7. (3p) Lös följande rekursionsekvation:

$$\begin{cases} a_0 = 1 \\ a_n = 3a_{n-1} + 4^n \quad \text{för } n = 1, 2, \dots \end{cases}$$

8. (3p) I beviset av följande sats om absorptionssannolikheter för markovkedjor har tre likheter markerats med (1),(2) och (3). Din uppgift är att förklara varför var och en av likheterna gäller. Om du hänvisar till en definition ska du presentera definitionen. Om du hänvisar till en sats ska du presentera den satsen. Det betyder att du ska tala om under vilka förutsättningar satsen gäller och kommentera varför förutsättningarna för den aktuella satsen är uppfyllda.

SATS

Betrakta en absorberande markovkedja med tillståndsrum $T = \{1, \dots, m\}$ och övergångsmatrix P . Låt a vara ett (bland flera) absorberande tillstånd för markovkedjan och låt

$$q_j = P(\text{kedjan absorberas i tillstånd } a \text{ då man startar i tillstånd } j)$$

Då är

$$q_j = \begin{cases} 1 & \text{om } j = a \\ 0 & \text{om } j \text{ är absorberande, men } j \neq a \\ \sum_{k=1}^m P_{jk} q_k & \text{annars} \end{cases}$$

BEVIS

Att $q_a = 1$ och att $q_j = 0$, om j är absorberande, men $j \neq a$, är uppenbart. Låt A vara händelsen att kedjan absorberas i tillstånd a . Om j ej är absorberande får vi:

$$\begin{aligned}
 q_j &= P(A|X_0 = j) \\
 &\stackrel{(1)}{=} \frac{P(A \cap X_0 = j)}{P(X_0 = j)} \\
 &\stackrel{(2)}{=} \frac{\sum_{k=1}^m P(A \cap X_0 = j \cap X_1 = k)}{P(X_0 = j)} \\
 &= \sum_{k=1}^m P(A|X_0 = j \cap X_1 = k) \frac{P(X_0 = j \cap X_1 = k)}{P(X_0 = j)} \\
 &= \sum_{k=1}^m P(A|X_0 = j \cap X_1 = k) P(X_1 = k|X_0 = j) \\
 &\stackrel{(3)}{=} \sum_{k=1}^m P(A|X_1 = k) P(X_1 = k|X_0 = j) \\
 &= \sum_{k=1}^m q_k P_{jk}
 \end{aligned}$$

9. (2p) Kal och Ada är två forskare med en del kunskap om Statistik. Kal studerar Göteborgares utlägg på julklappar i förhållande till deras inkomst. Ada studerar temperaturen mitt på dagen under november månad i Göteborg. Båda använder en enkel linjär regressionsmodell.

Kal definierar

X = månadsinkomsten för en göteborgare.

Y = göteborgarens utlägg på julklappar.

Nedan presenteras en del av Kals data.

x = inkomst	y = utlägg på julklappar
15000	1600
15000	1700
15000	2000
20000	2000
20000	2800
20000	2900
30000	2800
30000	3500
30000	5000
40000	3200
40000	6000
40000	7000
...	...

Ada definierar

X = datum i november.

Y = temperatur klockan 12.

Nedan presenteras en del av Adas data

x= datum i november	y = temperatur
1	5.8
2	5.8
3	5.7
4	5.3
5	4.9
6	4.9
7	4.8
8	5.3
9	5.4
10	5.2
11	5.2
12	5.1
13	4.6
14	4.5
15	4.3
16	4.4
17	4.2
...	...

Det var ett tag sedan Kal och Ada tog kursen i statistik. De behöver lite hjälp.

- Förklara för Kal varför en enkel linjär regressionsmodell inte verkar lämplig.
- Förklara för Ada varför en enkel linjär regressionsmodell inte verkar lämplig.

Lycka till!