

Eksamen i Statistik 2. år

Skriftlig prøve (4 timer)

19. maj 2009 kl. 9.00–13.00

Eksamenssættet er på 3 sider.

Alle skriftlige hjælpemidler samt lommeregner er tilladt.

Vægtfordeling: Opgaverne vægtes ens.

Opgave 1

Lad X være en reel stokastisk variabel med sandsynlighedsfunktion

$$p(x) = \begin{cases} \frac{x}{6} & \text{for } x \in \{1, 2, 3\} \\ 0 & \text{ellers.} \end{cases}$$

- (a) Udregn $P(X \leq 2)$, EX og $\text{var}(X)$.
- (b) Lad X_1 og X_2 være stokastisk uafhængige med samme fordeling som X . Opskriv sandsynlighedsfunktionen for $X_1 - X_2$, samt den betingede sandsynlighed for at $X_1 = 3$, givet at $X_1 - X_2 = 0$.
- (c) Lad $X_1, X_2, \dots, X_{100}$ være stokastisk uafhængige med samme fordeling som X . Hvad er, ifølge den centrale grænseværdisætning, den approksimative fordeling af $S = X_1 + X_2 + \dots + X_{100}$? Benyt dette til en approksimativ angivelse af 2.5%- og 97.5%-fraktilerne i S 's fordeling.

Opgave 2

Lad X være Γ -fordelt med formparameter 2, altså med tæthed

$$p(x) = \begin{cases} xe^{-x} & \text{for } x > 0 \\ 0 & \text{ellers.} \end{cases}$$

- (a) Angiv middelværdi og varians for $Y = 3 + 5X$.
- (b) Udregn tætheden for $Z = \exp(-X)$.
- (c) Lad X_1 og X_2 være uafhængige med samme fordeling som X . Angiv tætheden for $X_1 + X_2$.

Opgave 3

I nedenstående tabel er 328 kvindelige studerende ved et amerikansk universitet klassificeret efter deres øjenfarve (brun eller ikke brun) og hårfarve (sort, brun, rød eller blond).

	Brune øjne	Ikke brune øjne	I alt
1. Sort hår	41	11	52
2. Brunt hår	110	48	158
3. Rødt hår	23	14	37
4. Blondt hår	9	72	81
I alt	183	145	328

Vi betragter den binomialfordelingsmodel der fremkommer, når det antages at en tilfældig kvindelig studerende med hårfarve $h \in \{1, 2, 3, 4\}$ har brune øjne med sandsynlighed p_h .

(a) Estimer de fire sandsynlighedsparametre i denne model, med angivelse af 95% konfidensgrænser.

(b) Det er tydeligt, at vi ikke kan få godkendt hypotesen $p_1 = p_2 = p_3 = p_4$, fordi der er så få blandt de blonde som har brune øjne. Men kan det antages, at sorthårede, brunhårede og rødhårede har samme sandsynlighed for at være brunøjede?

(c) De tilsvarende tal for 264 mandlige studerende så sådan ud:

	Brune øjne	Ikke brune øjne	I alt
1. Sort hår	42	14	56
2. Brunt hår	63	65	128
3. Rødt hår	17	17	34
4. Blondt hår	8	38	46
I alt	130	134	264

Vi ser nu helt bort fra øjenfarven og betragter kun rækkesummerne i de to tabeller ovenfor. Opstil en model, hvor det har mening at spørge om fordelingen af hårfarven er den samme for de to køn, og foretag et test for denne hypotese.

Opgave 4

For 20 varegrupper, som forhandles i et supermarked, er det forudsagte salg x og det faktiske salg y på en bestemt dag (i antal stk.) registreret.

X	Y	X	Y	X	Y	X	Y
19	21	264	256	595	567	52	48
61	90	29	25	61	75	25	31
66	91	35	46	12	19	11	9
13	12	33	22	353	365	118	123
54	50	15	16	173	183	196	218

(a) Estimer parametrene i en simpel regressionsmodel med y som respons og x som forklarende variabel. For både afskæringens og hældningens vedkommende ønskes angivelse af 95% sikkerhedsgrænser. Ved udregningerne kan følgende mellemregningsstørrelser benyttes:

$$S_x = 2185 \quad S_y = 2267 \\ SS_x = 652817 \quad SS_y = 648671 \quad SP_{xy} = 648920$$

(b) Foretag ved hjælp af estimatorer og sikkerhedsgrænser fra spørgsmål (a) en vurdering af, om hypotesen $(\alpha, \beta) = (0, 1)$ kan accepteres (et egentligt test kræves ikke). Forklar betydningen af denne hypotese i den konkrete problemstilling. Her underforstås naturligvis, at α som sædvanlig betegner afskæringen, medens β er hældningen.

(c) Hvis man tegner y 'er mod x 'er (se nedenfor) ser man en svag tendens til, at punkterne grupperer sig om en konkav funktion snarere end en ret linie. Om dette er udtryk for en reel afvigelse fra regressionsmodellen kan man undersøge ved at tilføje et kvadratisk led, så modellen kommer til at se sådan ud:

$$E y_i = \mu_i = \alpha + \beta x_i + \gamma x_i^2.$$

I denne model fås $\hat{\gamma} = -0.0002632$, og estimatet for standardafvigelsen for $\hat{\gamma}$ er angivet til 0.00009498. Foretag et test for hypotesen $\gamma = 0$, og forklar konklusionen.

