

Oblig 3c

Gormery K. Wanjiru

13. april 2024

Innhold

1	(15%) kap. 17: oppgave 1.c	5
2	(15%) kap. 17: oppgave 1.d	5
3	Terningdropp-oppgaven: (Totalt 50%)	5
3.1	(5%) Tegn et diagram med samtlige datapunkter, og legg på den lineære regresjonslinjen.	5
3.2	(15%) Bruk nøytrale prior hyperparametre, og finn posterior og prediktive sannsynlighetsfordelinger, det vil si, sannsynlighetsfordelinger for τ , b , $y(x)$ og $Y^+(x)$	5
3.3	(5%) Finn et 80% kredibilitetsintervall (intervallestimat) for stigningstallet b	5
3.4	(5%) Finn et 80% kredibilitetsintervall (intervallestimat) for standardavviket σ . (Hint: Bruk verdiene fra τ og regn om ved å bruke at $\tau = \frac{1}{\sigma^2}$)	5
3.5	(5%) Finn et 80% kredibilitetsintervall (intervallestimat) for $y(x)$	5
3.6	(5%) 80% intervallestimatet for $y(x)$ er funksjoner av x , og en kurve over, og en under regresjonslinjen. Plott disse kurvene inn sammen med regresjonslinjen.	5
3.7	(5%) Finn verdien $R^2 = \frac{SS_y - SS_e}{SS_y}$. Dette tallet forteller hvor stor del av variasjonen i y som kan forklares av linja $y = a + bx$. For de av dere som bruker dataverktøy for å finne dette: angi hvordan dere fant det.	5
3.8	(5%) Finn R^2 for regresjonen mellom z (utfall på terningen) og x (dropphøyde). Kommenter hva forskjellen mellom R^2 for y og R^2 for z sier oss.	5
4	(Totalt 20%) Følgende R-kode vil plukke ut et utvalg av observasjonene.	5
4.1	(5%) Kjør 50 runder, og bruk $N = 15$. For hver runde, gjør oppgave 3a, men tegn regresjonslinjene sammen, i samme graf. Hva ser du?	6
4.2	(5%) Kjør en runde med N henholdsvis lik 5, 15, 50 og 200. For hver runde, gjør oppgavene 3c og 3d. Hva ser du?	6

4.3	(10%) Kjør en runde med N henholdsvis lik 5, 15, 50 og 200. For hver runde, gjør oppgaven 3f. Tegnes i hvert sitt diagram. Hva ser du?	6
Vedlegg		7
	Vedlegg A	7

- 1 (15%) kap. 17: oppgave 1.c
- 2 (15%) kap. 17: oppgave 1.d
- 3 Terningdropp-oppgaven: (Totalt 50%)
 - 3.1 (5%) Tegn et diagram med samtlige datapunkter, og legg på den lineære regresjonslinjen.
 - 3.2 (15%) Bruk nøytrale prior hyperparametre, og finn posterior og prediktive sannsynlighetsfordelinger, det vil si, sannsynlighetsfordelinger for τ , b , $y(x)$ og $Y^+(x)$.
 - 3.3 (5%) Finn et 80% kredibilitetsintervall (intervallestimat) for stigningstallet b .
 - 3.4 (5%) Finn et 80% kredibilitetsintervall (intervallestimat) for standardavviket σ . (Hint: Bruk verdiene fra τ og regn om ved å bruke at $\tau = \frac{1}{\sigma^2}$)
 - 3.5 (5%) Finn et 80% kredibilitetsintervall (intervallestimat) for $y(x)$.
 - 3.6 (5%) 80% intervallestimatet for $y(x)$ er funksjoner av x , og en kurve over, og en under regresjonslinjen. Plott disse kurvene inn sammen med regresjonslinjen.
 - 3.7 (5%) Finn verdien $R^2 = \frac{SS_y - SS_e}{SS_y}$. Dette tallet forteller hvor stor del av variasjonen i y som kan forklares av linja $y = a + bx$. For de av dere som bruker data-verktøy for å finne dette: angi hvordan dere fant det.
 - 3.8 (5%) Finn R^2 for regresjonen mellom z (utfall på terningen) og x (dropphøyde). Kommenter hva forskjellen mellom R^2 for \tilde{y} og R^2 for z sier oss.
- 4 (Totalt 20%) Følgende R-kode vil plukke ut et utvalg av observasjonene.

R kode

```
N = 20 # (Eksempel; se oppgavene)
utvalg = sort(sample(1:antall_rader,N)) # Sortering er ikke nødvendig
utvalg # men du får da se hvilke rader som er plukket ut
ny_dropp_df = dropp_df[utvalg,] # Dette er kjernen; plukker ut radene
rownames(ny_dropp_df)=1:N # Lurt hvis du skal kjøre for-løkke.
ny_dropp_df # Ikke nødvendig, men du får se den nye data-rammen.
```

- 4.1 (5%) Kjør 50 runder, og bruk $N = 15$. For hver runde, gjør oppgave 3a, men tegn regresjonslinjene sammen, i samme graf. Hva ser du?
- 4.2 (5%) Kjør en runde med N henholdsvis lik 5, 15, 50 og 200. For hver runde, gjør oppgavene 3c og 3d. Hva ser du?
- 4.3 (10%) Kjør en runde med N henholdsvis lik 5, 15, 50 og 200. For hver runde, gjør oppgaven 3f. Tegnes i hvert sitt diagram. Hva ser du?

Vedlegg

Vedlegg A

Referanser

- [1] <https://tma4245.math.ntnu.no/viktige-diskrete-fordelinger/poissonprosess-og-poissonfordeling> *NTNU*