Oblig 3b

Levering: 1 PDF, i rett mappe på Canvas. Lever eventuell **R**/**MatLab**/**Wolfram**-kode som kildefil *i tillegg*.
Førstefrist: 23. mars, 18:00
Sistefrist: 6. april, 18:00
Godkjent: 40% + (antall i gruppa) $*$ 10%

1. (10%) Begreper
   (a) Hva er en Poisson-prosess?
   (b) Hva er *prediktiv* fordeling?
   (c) Hva er sammenhengen mellom gamma-fordeling og gamma-gamma-fordeling, og hva er forskjellen på de to?
2. (10%) Kapittel 13: oppgave 13d
3. (10%) Kapittel 13: oppgave 14c
4. **De tre eksperimentene:**
   (a) **Bernoulli (30%):** Du skal her ta utgangspunkt i de 4 sannsynlighetsfordelingene du fikk og tegnet for $n = 10, 100, 1000$ og 10000 i oblig 3a.
      i. Intervallestimater: Regn ut 90% kredibilitetsintervall for parameteren π for alle 4 n-verdier. Fremstill de 4 intervallene sammen på en måte som viser utviklingen etter hvert som du samler mer informasjon. Enkle linjer tegnet for hånd med forskjellige farger for å markere områdene til intervallene er tilstrekkelig.
      ii. Punktestimater: For $n = 100$ og $n = 10000$, regn ut de tre punktestimatene for β-fordelte variable for $\beta_{(a_{100}, b_{100})}$ og for $\beta_{(a_{10000}, b_{10000})}$. Tegn de to grafene hver for seg sammen med de tre estimatene sine. (for ilustrasjonene her anbefales regneverktøy som **R**, **MatLab** eller **Wolfram** sterkt fremfor frihåndtegning!).
      iii. Du skal avgjøre hypotesen H_1: $\pi > 0.5$ med signifikans $\alpha = 0.1$. Regn ut $P(H_0)$ for alle de fire *posterior*-fordelingene (for $n = 10, 100, 1000$ og 10000), og avgjør hypotesen for hver av dem.
      iv. (BONUS) Finn et tall p_0 slik at du endrer konklusjon på hypotesen H_1: $\pi > p_0$ etter hvert som n blir større.
      v. Sammenligning: Ta deres egen $\pi \sim \beta_{(a_{100}, b_{100})}$ og bytt verdier med en annen gruppe. Kall den andre gruppens variabel $p \sim \beta_{(A_{100}, B_{100})}$. Finn sannsynlighetene $P(\pi \geq p)$ og $P(p \geq \pi)$, og avgjør hypotesen H_1 at deres π er høyere enn den andre gruppas p med signifikans $\alpha = 0.2$.

   (b) **Poisson (20%):** (Bruk datasettet *discoveries*, som i 3a.)

i. Intervallestimater: Regn ut 20%, 40%, 60%, 80%, 90% og 95%, kredibilitetsintervall for parameteren λ for siste *posterior*-stadie av undersøkelsen. Fremstill de 6 intervallene sammen på en måte som viser utviklingen etter hvert som prosenten øker. Bruk gjerne konfidens-kurver. (for ilustrasjonene anbefales regneverktøy som **R**, **MatLab** eller **Wolfram** sterkt fremfor frihåndtegning!).

ii. Punktestimater: Regn ut de tre punktestimatene for den første av de tre posteriorene for λ.

iii. Du skal avgjøre hypotesen H_1: $\lambda > 2.5$ med signifikans $\alpha = 0.1$ for den siste av *posterior*-fordelingene dine.

iv. Sammenligning: Sammenlign din siste posterior for λ, raten for viktige oppdagelser per år, med λ_2, raten for viktige literære verk per år. Bruk $\lambda_2 \sim \gamma_{(147,50)}$. Avgjør hypotesen H_1: $\lambda > \lambda_2$ med signifikans $\alpha = 0.07$.

(c) **Gaussisk (20%):** I denne oppgaven skal dere bruke de *posterior* sannsynlighetsfordelingene dere fikk i oblig 3a. Velg to seigmanntyper (kall dem A og B), og gjør følgende:

i. Intervallestimater: Tegn konfidenskurvene for μ og for X_+ for seigmanntype A (for ilustrasjonene anbefales regneverktøy som **R**, **MatLab** eller **Wolfram** sterkt fremfor frihåndtegning!). Markér av på grafene 40%, 80%, og 90%, og les av manuelt fra grafene (altså øyemål, ikke beregning) hva de 40%, 80%, og 90% intervallestimatene for μ og for X_+ er.

ii. Si med ord hvorfor du tror det ikke ble noen oppgave om punktestimater for μ og X_+.

iii. Sammenligning 1: Test hypotesen H_1: $\mu_A > \mu_B$, med signifikans $\alpha = 0.042$

iv. Sammenligning 2: Test hypotesen H_1: $\tau_A > \tau_B$, med signifikans $\alpha = 0.042$