TMA4245 – Parallell 1

Geir-Arne Fuglstad, IMF

10. februar 2025

Gammafordeling, eksponentialfordeling og kjikvadratfordeling

Definisjon:
Den stokastiske variablen X følger en **gammafordeling** med **parametere** $\alpha, \beta > 0$ hvis sannsynlighetstettheten er gitt ved

$$f(x; \alpha, \beta) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} \exp\left(-\frac{x}{\beta}\right), \qquad x > 0.$$

Vi skriver f.eks. $X \sim \text{Gamma}(\alpha, \beta)$.

Definisjon:
Det finnes to veldig viktige spesialtilfeller med egne navn:

- **Eksponentialfordelingen** ($\alpha = 1$) med **skaleringsparameter** $\beta > 0$. Dvs. $X \sim \text{Exp}(\beta)$,

$$f(x; \beta) = \frac{1}{\beta} \exp\left(-\frac{x}{\beta}\right), \qquad x > 0.$$

- **Kjikvadratfordelingen** ($\alpha = \nu/2$ og $\beta = 2$) med $\nu > 0$ **frihetsgrader**. Dvs. $X \sim \chi^2_\nu$,

$$f(x; \nu) = \frac{1}{2^{\nu/2} \Gamma(\nu/2)} x^{\nu/2-1} \exp\left(-\frac{x}{2}\right), \qquad x > 0.$$

ADVARSEL om parametrisering

Alternativ 1 (Formelsamling):

$$f(x;\alpha,\beta) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} \exp\left(-\frac{x}{\beta}\right), \qquad x > 0,$$

her kalles parameterene ofte α “shape parameter” og β “scale parameter”.

Alternativ 2 (Koblet til poissonprosessen):

$$f(x;\alpha,\lambda) = \frac{\lambda^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\lambda x), \qquad x > 0,$$

her kalles parameterene ofte α “shape parameter” og λ “rate parameter”.

Begge disse er ekstremt vanlige. Man må sjekke nøye hva forfatter mener i enhver tekst man leser og kode man bruker. Sammenhengen $\lambda = 1/\beta$ er enkel, men det er kritisk at det blir riktig!

Egenskaper for gammafordelingen

Forventningsverdi og varians:

$$\mathrm{E}[X] = \alpha\beta \quad \text{og} \quad \mathrm{Var}[X] = \alpha\beta^2$$

gitt i formelsamling, men vær OBS på parametrisering.

Punktsannsynlighet og kumulativ fordeling:

Punktsannsynlighet gitt i formelsamling. Kumulativ fordeling er ikke tabellert i formelsamlingen, men dere kan finne "kritiske verdier" i kjikvadratfordelingen. Dvs $\chi^2_{\alpha,\nu}$ slik at

$$P\left(X > \chi^2_{\alpha,\nu}\right) = \alpha$$

når $X \sim \chi^2_\nu$. Disse trenger dere senere i kurset.

NB: "Memoryless"

Eksponentialfordelingen er den eneste kontinuerlige fordelingen på $(0,\infty)$ som oppfyller

$$\mathrm{P}(X > s + t | X > t) = \mathrm{P}(X > s), \qquad \text{for alle } t, s \geq 0.$$

På lignende måte er den geometriske fordelingen den eneste diskret fordelingen (på $\mathbb{N}$) slik at

$$\mathrm{P}(X > m + n | X > n) = \mathrm{P}(X > m), \qquad \text{for alle } m, n \in \mathbb{N}$$

Python

Sjekk referansesider for scipy.stats:
https://docs.scipy.org/doc/scipy/reference/stats.html

Søk på siden etter den fordelingen du ønsker, og les beskrivelsen. Det finnes funksjoner for cdf, kvantiler, pdf/pmf, og simulering.

F.eks. gammafordelingen:
https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.gamma.html#scipy.stats.gamma

(NB: Kan også sjekke numpy.random, men denne er ikke i fortsatt utvikling: https://numpy.org/doc/stable/reference/random/legacy.html)

t-fordelingen

Definisjon:
En stokastisk variabel X er ***t*-fordelt** med $\nu > 0$ **frihetsgrader** hvis sannsynlighetstettheten er

$$f(x;\nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi\nu}}\left(1+\frac{x^2}{\nu}\right)^{-(\nu+1)/2}, \qquad x \in \mathbb{R}.$$

Kan skrives f.eks. $X \sim t_\nu$.

Denne fordelingen oppstår i uke 9/10 når vi har to uavhengige variabler $Z \sim \mathrm{N}(0,1)$ og $V \sim \chi^2_\nu$ og trenger fordelingen til

$$T = \frac{Z}{\sqrt{V/\nu}}.$$

Jeg planlegger å snakke mer om fordelingen når vi kommer dit.

Transformasjoner av én stokastisk variabel

For en stokastisk variabel X og strengt monoton funksjon g:

1. Hvis X har punktsannsynlighet
$$f_X(x), \qquad x = x_1, x_2, \dots,$$
har $Y = g(X)$ punktsannsynlighet
$$f_Y(y) = f_X\big(g^{-1}(y)\big), \qquad y = g(x_1), g(x_2), \dots$$
2. Hvis X har sannsynlighetstetthet
$$f_X(x), \qquad a < x < b,$$
har $Y = g(X)$ sannsynlighetstetthet
$$f_Y(y) = f_X\big(g^{-1}(y)\big)\left|\frac{d}{dy}g^{-1}(y)\right|, \qquad g(a) < y < g(b).$$

Mer om transformasjon

Ikke strengt monoton
Hvis g ikke er strengt monoton må man dele opp i strengt monotone deler og behandle hver del separat. F.eks. $Y = X^2$ krever at vi behandler $g_1^{-1}(y) = \sqrt{y}$ og $g_2^{-1}(y) = -\sqrt{y}$ separat. Se introvideo.

Gauss' feilforplantningslov:
Anta uavhengige stokastiske variabler $X_1, X_2, \dots, X_n$ med forventningsverdier $\mu_i, i = 1, \dots, n$, og varianser $\sigma_i^2, i = 1, \dots, n$.

For en deriverbar funksjon $g: \mathbb{R}^n \to \mathbb{R}$, har vi for den stokastiske variabelen $Y = g(X_1, X_2, \dots, X_n)$ tilnærmingene

$$\mathrm{E}[Y] = \mathrm{E}[g(X_1, \dots, X_n)] \approx g(\mu_1, \dots, \mu_n)$$

$$\mathrm{Var}[Y] = \mathrm{Var}[g(X_1, \dots, X_n)] \approx \sum_{k=1}^{n} \left(\frac{\partial g}{\partial x_k}(\mu_1, \dots, \mu_n) \right)^2 \sigma_k^2$$