

STK4021 - Applied Bayesian Analysis

Useful distributions

Odd Kolbjørnsen)

Fall 2026

- The Gamma distribution is defined by:

$$p(\theta | \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta}, \quad \theta > 0,$$

where $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$ is the Gamma function.

- For integer α

$$\Gamma(\alpha) = (n-1)!$$

- $\mathbb{E}[\theta] = \frac{\alpha}{\beta},$
- $\text{Var}(\theta) = \frac{\alpha}{\beta^2}$
- Special cases:
 - $\alpha = 1$ is the exponential distribution
 - $(\alpha, \beta) = (\nu/2, 1/2)$ is the χ^2 distribution with ν degrees of freedom.

- The Inverse-Gamma distribution is defined by:

$$p(\theta \mid \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{-(\alpha+1)} \exp\left(-\frac{\beta}{\theta}\right), \quad \theta > 0,$$

- $\mathbb{E}[\theta] = \frac{\beta}{\alpha-1}, \quad \alpha > 1$
- $\text{Var}(X) = \frac{\beta^2}{(\alpha-1)^2(\alpha-2)}, \quad \alpha > 2.$
- If X has the Gamma distribution the $1/X$ has the inverse-Gamma distribution

- The Beta distribution is defined by

$$p(\theta|\alpha, \beta) = \frac{1}{B(\alpha, \beta)} \theta^{\alpha-1} (1 - \theta)^{\beta-1}, \quad \theta \in [0, 1]$$

where $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$

- For α, β integers,

$$B(\alpha, \beta) = \frac{(\alpha - 1)!(\beta - 1)!}{(\alpha + \beta - 1)!}$$

- $\mathbb{E}[X] = \frac{\alpha}{\alpha + \beta}$
- $\text{Var}(X) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$
- For $\alpha = \beta = 1$ we obtain the **uniform** distribution

- The Dirichlet distribution is defined by

$$p(\theta \mid \alpha) = \frac{\Gamma(\alpha_0)}{\prod_{i=1}^K \Gamma(\alpha_i)} \prod_{i=1}^K \theta_i^{\alpha_i-1},$$

$$\alpha_0 = \sum_{i=1}^K \alpha_i, \quad \theta_i \geq 0, \quad \sum_{i=1}^K \theta_i = 1.$$

- $\mathbb{E}[\theta_i \mid \alpha] = \frac{\alpha_i}{\alpha_0}.$
- $\text{Var}(\theta_i \mid \alpha) = \frac{\alpha_i(\alpha_0 - \alpha_i)}{\alpha_0^2(\alpha_0 + 1)}.$
- $\text{Cov}(\theta_i, \theta_j \mid \alpha) = -\frac{\alpha_i \alpha_j}{\alpha_0^2(\alpha_0 + 1)}, \quad i \neq j.$
- Marginal $\theta_i \sim \text{Beta}(\alpha_i, \alpha_0 - \alpha_i), \quad \alpha_0 = \sum_{j=1}^K \alpha_j$
- Summing components entails to sum parameters, e.g.
 $(\theta_1, \theta_2, \theta_3, \theta_4) \sim \text{Dirichlet}(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$
 $\implies (\theta_1 + \theta_2, \theta_3, \theta_4) \sim \text{Dirichlet}(\alpha_1 + \alpha_2, \alpha_3, \alpha_4)$

- The Binomial distribution is defined by

$$p(x | \theta) = \binom{n}{x} \theta^x (1 - \theta)^{n-x}, \quad x = 0, 1, \dots, n.$$

where $\binom{n}{x} = \frac{n!}{x!(n-x)!}$

- $\mathbb{E}[X | \theta] = n\theta$
- $\text{Var}(X | \theta) = n\theta(1 - \theta)$
- $n = 1$ gives a Bernoulli trial

- The distribution is defined by

$$p(x | \theta) = (1 - \theta)^x \theta, \quad x = 0, 1, 2, \dots$$

- $\mathbb{E}[X | \theta] = \frac{1-\theta}{\theta},$
- $\text{Var}(X | \theta) = \frac{1-\theta}{\theta^2}.$
- Number of failures until the first success
- Alternative:
Number of trials until the first success: $Y = X + 1$

$$p(y | \theta) = (1 - \theta)^{y-1} \theta, \quad y = 1, 2, \dots$$

- $\mathbb{E}[Y] = \frac{1}{\theta}$
- $\text{Var}(Y) = \frac{1-\theta}{\theta^2}$

- The Poisson distribution is defined by

$$p(x | \theta) = \frac{\theta^x e^{-\theta}}{x!}, \quad x = 0, 1, 2, \dots, \quad \theta > 0.$$

- $\mathbb{E}[X | \theta] = \theta$,
- $\text{Var}(X | \theta) = \theta$.

The negative binomial distribution

- The negative binomial distribution is defined by

$$p(y \mid \alpha, \beta) = \frac{\Gamma(y + \alpha)}{\Gamma(\alpha) y!} \left(\frac{\beta}{\beta + 1} \right)^\alpha \left(\frac{1}{\beta + 1} \right)^y, \quad y = 0, 1, 2, \dots$$

- For α integer $\frac{\Gamma(y + \alpha)}{\Gamma(\alpha) y!} = \binom{y + \alpha - 1}{\alpha - 1} = \binom{y + \alpha - 1}{y}$.



- The Normal distribution is defined by

$$p(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{(x - \mu)^2}{2\sigma^2} \right], \quad x \in \mathbb{R}.$$

- $\mathbb{E}[X \mid \mu, \sigma^2] = \mu,$
- $\text{Var}(X \mid \mu, \sigma^2) = \sigma^2$
- Sum of two independent Normal variables :

$$X \sim \mathcal{N}(\mu_X, \sigma_X^2) \text{ and } Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$$

$$\Rightarrow X + Y \sim \mathcal{N}(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2).$$

The Multivariate normal distribution

- The Multivariate normal distribution $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is defined by

$$p(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{p/2} |\boldsymbol{\Sigma}|^{1/2}} \exp \left[-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right], \quad \mathbf{x} \in \mathbb{R}^p.$$

- $\mathbb{E}[\mathbf{X}] = \boldsymbol{\mu}$, and $\text{Cov}(\mathbf{X}) = \boldsymbol{\Sigma}$.
- Linear transform:

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b} \implies \mathbf{Y} \sim \mathcal{N}_q(\mathbf{A}\boldsymbol{\mu} + \mathbf{b}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^\top)$$

- Conditional distribution:

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix} \right).$$

- Then

$$\mathbf{x}_1 \mid \mathbf{x}_2 \sim \mathcal{N}(\boldsymbol{\mu}_{1|2}, \boldsymbol{\Sigma}_{1|2}),$$

- Where

$$\boldsymbol{\mu}_{1|2} = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2), \text{ and } \boldsymbol{\Sigma}_{1|2} = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}.$$

The Multivariate normal distribution, precision

- The Multivariate normal distribution is defined by:

$$p(\mathbf{x} \mid \boldsymbol{\eta}, \mathbf{Q}) \propto |\mathbf{Q}|^{1/2} \exp \left[-\frac{1}{2} \mathbf{x}^T \mathbf{Q} \mathbf{x} + \mathbf{x}^T \boldsymbol{\eta} - \frac{1}{2} \boldsymbol{\eta}^T \mathbf{Q}^{-1} \boldsymbol{\eta} \right], \quad \mathbf{x} \in \mathbb{R}^p.$$

- $\mathbb{E}[\mathbf{X}] = \mathbf{Q}^{-1} \boldsymbol{\eta}$, $\text{Cov}(\mathbf{X}) = \mathbf{Q}^{-1}$ with $(\boldsymbol{\eta} = \mathbf{Q} \boldsymbol{\mu}, \mathbf{Q} = \boldsymbol{\Sigma}^{-1})$
- Linear transform:

$$\mathbf{Y} = \mathbf{A} \mathbf{X} + \mathbf{b} \implies \mathbf{Y} \sim \mathcal{N}_q(\mathbf{A} \boldsymbol{\mu} + \mathbf{b}, \mathbf{A} \mathbf{Q}^{-1} \mathbf{A}^T)$$

- Conditional distribution:

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix}^{-1} \right),$$

- Then

$$\mathbf{x}_1 \mid \mathbf{x}_2 \sim \mathcal{N}(\boldsymbol{\mu}_{1|2}, \mathbf{Q}_{11}^{-1}),$$

- Where

$$\boldsymbol{\mu}_{1|2} = \boldsymbol{\mu}_1 - \mathbf{Q}_{11}^{-1} \mathbf{Q}_{12} (\mathbf{x}_2 - \boldsymbol{\mu}_2). \text{ Note: } \boldsymbol{\Sigma}_{1|2} = \mathbf{Q}_{11}^{-1}.$$