SDS 383C: Statistical Modeling I Fall 2022, Module V

Abhra Sarkar

Department of Statistics and Data Sciences The University of Texas at Austin

"All models are wrong, but some are useful."- George E. P. Box

$$y_1, \ldots, y_n \stackrel{iid}{\sim} \text{Normal}(\mu, \sigma^2)$$
 with μ, σ^2 both unknown

- Normal Likelihood: $p(\mathbf{y}_{1:n} \mid \mu, \sigma^2) \propto (\sigma^2)^{-\frac{n}{2}} \exp \left[-\frac{1}{2\sigma^2} \left\{ (n-1)s^2 + n(\overline{y} \mu)^2 \right\} \right]$
- ▶ Normal-Inverse-Gamma Prior: $(\mu, \sigma^2) \sim \text{NIG}(\mu_0, \sigma_0^2/\kappa_0, \nu_0, \sigma_0^2)$

▶ Normal-Inverse-Gamma Posterior:

$$\begin{split} & \rho(\mu, \sigma^2 \mid \mathbf{y}_{1:n}) \\ & \propto \left(\sigma^2\right)^{-\left\{\frac{(\nu_0 + n)}{2} + 1 + \frac{1}{2}\right\}} \exp\left[-\frac{1}{2\sigma^2} \left\{\nu_0 \sigma_0^2 + (n - 1)s^2 + \frac{n\kappa_0}{(n + \kappa_0)} (\overline{y} - \mu_0)^2 + (\kappa_0 + n)(\mu - \mu_n)^2\right\}\right] \\ & \equiv \text{NIG}\left(\mu_n, \sigma_n^2 / \kappa_n, \nu_n, \sigma_n^2\right), \quad \nu_n = (\nu_0 + n), \quad \kappa_n = (\kappa_0 + n), \quad \mu_n = (\kappa_0 \mu_0 + n\overline{y})/(\kappa_0 + n), \end{split}$$

$$\sigma_n^2 = \frac{1}{\nu_n} \left\{ \nu_0 \, \sigma_0^2 + (n-1) \, s^2 + \frac{n \kappa_0}{(n+\kappa_0)} (\overline{v} - \mu_0)^2 \right\} \tag{WhiteBoard}$$

2/18 2/71

$$y_1, \ldots, y_n \stackrel{iid}{\sim} \text{Normal}(\mu, \sigma^2)$$
 with μ, σ^2 both unknown

- Normal Likelihood: $p(\mathbf{y}_{1:n} \mid \mu, \sigma^2) \propto (\sigma^2)^{-\frac{n}{2}} \exp \left[-\frac{1}{2\sigma^2} \left\{ (n-1)s^2 + n(\overline{y} \mu)^2 \right\} \right]$
- ▶ Normal-Inverse-Gamma Prior: $(\mu, \sigma^2) \sim \text{NIG}(\mu_0, \sigma_0^2/\kappa_0, \nu_0, \sigma_0^2)$

$$\propto \left(\sigma^{2}\right)^{-\left\{\frac{(\nu_{0}+n)}{2}+1+\frac{1}{2}\right\}} \exp\left[-\frac{1}{2\sigma^{2}}\left\{\nu_{0}\sigma_{0}^{2}+(n-1)s^{2}+\frac{n\kappa_{0}}{(n+\kappa_{0})}(\overline{y}-\mu_{0})^{2}+(\kappa_{0}+n)(\mu-\mu_{n})^{2}\right]$$

$$\sigma_n^2 = \frac{1}{\nu_n} \left\{ \nu_0 \, \sigma_0^2 + (n-1)s^2 + \frac{n\kappa_0}{(n+\kappa_0)} (\overline{y} - \mu_0)^2 \right\} \tag{WhiteBoard}$$

2/18 3/71

$$y_1, \ldots, y_n \stackrel{iid}{\sim} \text{Normal}(\mu, \sigma^2)$$
 with μ, σ^2 both unknown

- Normal Likelihood: $p(\mathbf{y}_{1:n} \mid \mu, \sigma^2) \propto (\sigma^2)^{-\frac{n}{2}} \exp \left[-\frac{1}{2\sigma^2} \left\{ (n-1)s^2 + n(\overline{y} \mu)^2 \right\} \right]$
- Normal-Inverse-Gamma Prior: $(\mu, \sigma^2) \sim \text{NIG}(\mu_0, \sigma_0^2/\kappa_0, \nu_0, \sigma_0^2)$

$$\begin{split} &p(\mu,\sigma^2) = p(\sigma^2)p(\mu \mid \sigma^2) \\ &= \text{Inv-Ga}\left(\sigma^2 \mid \nu_0/2,\nu_0\sigma_0^2/2\right) \cdot \text{Normal}(\mu \mid \mu_0,\sigma^2/\kappa_0) \\ &= \frac{(\nu_0\sigma_0^2)^{\frac{\nu_0}{2}}}{\Gamma(\nu_0/2)\left(\sigma^2\right)^{(\frac{\nu_0}{2}+1)}} \exp\left(-\frac{\nu_0\sigma_0^2}{2\sigma^2}\right) \cdot \frac{\sqrt{\kappa_0}}{\sqrt{2\pi}\sigma} \exp\left[-\frac{\kappa_0}{2\sigma^2}(\mu-\mu_0)^2\right]^{\frac{1}{180}} \\ &\propto \left(\sigma^2\right)^{-\left(\frac{\nu_0}{2}+1+\frac{1}{2}\right)} \exp\left[-\frac{1}{2\sigma^2}\left\{\nu_0\sigma_0^2+\kappa_0(\mu-\mu_0)^2\right\}\right] \end{split}$$

▶ Normal-Inverse-Gamma Posterior:

$$p(\mu, \sigma^2 \mid \mathbf{y}_{1:n})$$

$$\propto \left(\sigma^{2}\right)^{-\left\{\frac{(\nu_{0}+n)}{2}+1+\frac{1}{2}\right\}} \exp\left[-\frac{1}{2\sigma^{2}}\left\{\nu_{0}\sigma_{0}^{2}+(n-1)s^{2}+\frac{n\kappa_{0}}{(n+\kappa_{0})}(\overline{y}-\mu_{0})^{2}+(\kappa_{0}+n)(\mu-\mu_{n})^{2}\right\}\right]$$

 $\equiv \operatorname{NIG}\left(\mu_n, \sigma_n^2/\kappa_n, \nu_n, \sigma_n^2\right), \quad \nu_n = (\nu_0 + n), \quad \kappa_n = (\kappa_0 + n), \quad \mu_n = (\kappa_0 \mu_0 + n\overline{y})/(\kappa_0 + n),$

$$\sigma_n^2 = \frac{1}{\nu_n} \left\{ \nu_0 \sigma_0^2 + (n-1)s^2 + \frac{n\kappa_0}{(n+\kappa_0)} (\overline{y} - \mu_0)^2 \right\} \tag{WhiteBoard}$$

2/18 4/71

$$y_1, \ldots, y_n \stackrel{iid}{\sim} \text{Normal}(\mu, \sigma^2)$$
 with μ, σ^2 both unknown

- Normal Likelihood: $p(\mathbf{y}_{1:n} \mid \mu, \sigma^2) \propto (\sigma^2)^{-\frac{n}{2}} \exp \left[-\frac{1}{2\sigma^2} \left\{ (n-1)s^2 + n(\overline{y} \mu)^2 \right\} \right]$
- Normal-Inverse-Gamma Prior: $(\mu, \sigma^2) \sim \text{NIG}(\mu_0, \sigma_0^2/\kappa_0, \nu_0, \sigma_0^2)$

$$\begin{split} & p(\mu,\sigma^2) = p(\sigma^2)p(\mu \mid \sigma^2) \\ & = \text{Inv-Ga}\left(\sigma^2 \mid \nu_0/2, \nu_0\sigma_0^2/2\right) \cdot \text{Normal}(\mu \mid \mu_0,\sigma^2/\kappa_0) \\ & = \frac{(\nu_0\sigma_0^2)^{\frac{\nu_0}{2}}}{\Gamma(\nu_0/2)(\sigma^2)^{(\frac{\nu_0}{2}+1)}} \exp\left(-\frac{\nu_0\sigma_0^2}{2\sigma^2}\right) \cdot \frac{\sqrt{\kappa_0}}{\sqrt{2\pi}\sigma} \exp\left[-\frac{\kappa_0}{2\sigma^2}(\mu-\mu_0)^2\right]^{\frac{1}{2}} \\ & \propto \left(\sigma^2\right)^{-\left(\frac{\nu_0}{2}+1+\frac{1}{2}\right)} \exp\left[-\frac{1}{2\sigma^2}\left\{\nu_0\sigma_0^2+\kappa_0(\mu-\mu_0)^2\right\}\right] \end{split}$$

► Normal-Inverse-Gamma Posterior:

$$\begin{split} &p(\mu,\sigma^2\mid\mathbf{y}_{1:n})\\ &\propto \left(\sigma^2\right)^{-\left\{\frac{(\nu_0+n)}{2}+1+\frac{1}{2}\right\}} \exp\left[-\frac{1}{2\sigma^2}\left\{\nu_0\sigma_0^2+(n-1)s^2+\frac{n\kappa_0}{(n+\kappa_0)}(\overline{y}-\mu_0)^2+(\kappa_0+n)(\mu-\mu_n)^2\right\}\right]\\ &\equiv \mathrm{NIG}\left(\mu_n,\sigma_n^2/\kappa_n,\nu_n,\sigma_n^2\right),\quad \nu_n=(\nu_0+n),\quad \kappa_n=(\kappa_0+n),\quad \mu_n=(\kappa_0\mu_0+n\overline{y})/(\kappa_0+n),\\ &\sigma_n^2=\frac{1}{\nu_n}\left\{\nu_0\sigma_0^2+(n-1)s^2+\frac{n\kappa_0}{(n+\kappa_0)}(\overline{y}-\mu_0)^2\right\} \end{split} \tag{WhiteBoard}$$

$$y_1, \dots, y_n \overset{iid}{\sim} \operatorname{Normal}(\mu, \sigma^2)$$
 with μ, σ^2 both unknown

- Normal Likelihood: $p(\mathbf{y}_{1:n} \mid \mu, \sigma^2) \propto (\sigma^2)^{-\frac{n}{2}} \exp \left[-\frac{1}{2\sigma^2} \left\{ (n-1)s^2 + n(\overline{y} \mu)^2 \right\} \right]$
- ▶ Normal-Inverse-Gamma Prior: $(\mu, \sigma^2) \sim \text{NIG}\left(\mu_0, \sigma_0^2 / \kappa_0, \nu_0, \sigma_0^2\right)$

$$\begin{split} p(\mu,\sigma^2) &= p(\sigma^2) p(\mu \mid \sigma^2) = \text{Inv-Ga}\left(\sigma^2 \mid \nu_0/2, \nu_0 \sigma_0^2/2\right) \cdot \text{Normal}(\mu \mid \mu_0, \sigma^2/\kappa_0) \\ &\propto \left(\sigma^2\right)^{-\left(\frac{\nu_0}{2} + 1 + \frac{1}{2}\right)} \exp\left[-\frac{1}{2\sigma^2} \left\{\nu_0 \sigma_0^2 + \kappa_0 (\mu - \mu_0)^2\right\}\right] \end{split}$$

► Normal-Inverse-Gamma Posterior:

$$\begin{split} & p(\mu,\sigma^2\mid\mathbf{y}_{1:n})\\ & \propto \left(\sigma^2\right)^{-\left\{\frac{(\nu_0+n)}{2}+1+\frac{1}{2}\right\}} \exp\left[-\frac{1}{2\sigma^2}\left\{\nu_0\sigma_0^2+(n-1)s^2+\frac{n\kappa_0}{(n+\kappa_0)}(\overline{y}-\mu_0)^2+(\kappa_0+n)(\mu-\mu_n)^2\right\}\right]\\ & \equiv \mathrm{NIG}\left(\mu_n,\sigma_n^2/\kappa_n,\nu_n,\sigma_n^2\right),\quad \nu_n=(\nu_0+n),\quad \kappa_n=(\kappa_0+n),\quad \mu_n=(\kappa_0\mu_0+n\overline{y})/(\kappa_0+n),\\ & \sigma_n^2=\frac{1}{\nu_n}\left\{\nu_0\sigma_0^2+(n-1)s^2+\frac{n\kappa_0}{(n+\kappa_0)}(\overline{y}-\mu_0)^2\right\} \end{split}$$

▶ A-priori and a-posteriori μ and σ^2 are dependent but uncorrelated.

3/18 6/71

$$y_1, \dots, y_n \overset{iid}{\sim} \operatorname{Normal}(\mu, \sigma^2)$$
 with μ, σ^2 both unknown

- Normal Likelihood: $p(\mathbf{y}_{1:n} \mid \mu, \sigma^2) \propto (\sigma^2)^{-\frac{n}{2}} \exp \left[-\frac{1}{2\sigma^2} \left\{ (n-1)s^2 + n(\overline{y} \mu)^2 \right\} \right]$
- ▶ Normal-Inverse-Gamma Prior: $(\mu, \sigma^2) \sim \text{NIG}\left(\mu_0, \sigma_0^2 / \kappa_0, \nu_0, \sigma_0^2\right)$

$$\begin{split} p(\mu,\sigma^2) &= p(\sigma^2) p(\mu \mid \sigma^2) = \text{Inv-Ga}\left(\sigma^2 \mid \nu_0/2, \nu_0 \sigma_0^2/2\right) \cdot \text{Normal}(\mu \mid \mu_0, \sigma^2/\kappa_0) \\ &\propto \left(\sigma^2\right)^{-\left(\frac{\nu_0}{2} + 1 + \frac{1}{2}\right)} \exp\left[-\frac{1}{2\sigma^2} \left\{\nu_0 \sigma_0^2 + \kappa_0 (\mu - \mu_0)^2\right\}\right] \end{split}$$

► Normal-Inverse-Gamma Posterior:

$$\begin{split} & p(\mu,\sigma^2\mid\mathbf{y}_{1:n})\\ & \propto \left(\sigma^2\right)^{-\left\{\frac{(\nu_0+n)}{2}+1+\frac{1}{2}\right\}} \exp\left[-\frac{1}{2\sigma^2}\left\{\nu_0\sigma_0^2+(n-1)s^2+\frac{n\kappa_0}{(n+\kappa_0)}(\overline{y}-\mu_0)^2+(\kappa_0+n)(\mu-\mu_n)^2\right\}\right]\\ & \equiv \mathrm{NIG}\left(\mu_n,\sigma_n^2/\kappa_n,\nu_n,\sigma_n^2\right),\quad \nu_n=(\nu_0+n),\quad \kappa_n=(\kappa_0+n),\quad \mu_n=(\kappa_0\mu_0+n\overline{y})/(\kappa_0+n),\\ & \sigma_n^2=\frac{1}{\nu_n}\left\{\nu_0\sigma_0^2+(n-1)s^2+\frac{n\kappa_0}{(n+\kappa_0)}(\overline{y}-\mu_0)^2\right\} \end{split}$$

ightharpoonup A-priori and a-posteriori μ and σ^2 are dependent but uncorrelated. HW.4.1a

3/18 7/71

Normal Model

Conditional Posteriors under the Conjugate Prior

$$y_1, \ldots, y_n \stackrel{iid}{\sim} \text{Normal}(\mu, \sigma^2)$$
 with μ, σ^2 both unknown

- Normal Likelihood: $p(\mathbf{y}_{1:n} \mid \mu, \sigma^2) \propto (\sigma^2)^{-\frac{n}{2}} \exp \left[-\frac{1}{2\sigma^2} \left\{ (n-1)s^2 + n(\overline{y} \mu)^2 \right\} \right]$
- ▶ Normal-Inverse-Gamma Prior: $(\mu, \sigma^2) \sim \text{NIG}\left(\mu_0, \sigma_0^2 / \kappa_0, \nu_0, \sigma_0^2\right)$

$$\begin{split} p(\mu,\sigma^2) &= p(\sigma^2) p(\mu \mid \sigma^2) = \text{Inv-Ga}\left(\sigma^2 \mid \nu_0/2,\nu_0\sigma_0^2/2\right) \cdot \text{Normal}(\mu \mid \mu_0,\sigma^2/\kappa_0) \\ &\propto \left(\sigma^2\right)^{-\left(\frac{\nu_0}{2}+1+\frac{1}{2}\right)} \exp\left[-\frac{1}{2\sigma^2}\left\{\nu_0\sigma_0^2 + \kappa_0(\mu-\mu_0)^2\right\}\right] \end{split}$$

▶ Normal-Inverse-Gamma Posterior:

$$\begin{split} & p(\mu,\sigma^2 \mid \mathbf{y}_{1:n}) \\ & \propto \left(\sigma^2\right)^{-\left\{\frac{(\nu_0+n)}{2}+1+\frac{1}{2}\right\}} \exp\left[-\frac{1}{2\sigma^2}\left\{\nu_0\sigma_0^2+(n-1)s^2+\frac{n\kappa_0}{(n+\kappa_0)}(\overline{y}-\mu_0)^2+(\kappa_0+n)(\mu-\mu_n)^2\right\}\right] \\ & \equiv \mathrm{NIG}\left(\mu_n,\sigma_n^2/\kappa_n,\nu_n,\sigma_n^2\right), \quad \nu_n=(\nu_0+n), \quad \kappa_n=(\kappa_0+n), \quad \mu_n=(\kappa_0\mu_0+n\overline{y})/(\kappa_0+n), \\ & \sigma_n^2=\frac{1}{\nu_n}\left\{\nu_0\sigma_0^2+(n-1)s^2+\frac{n\kappa_0}{(n+\kappa_0)}(\overline{y}-\mu_0)^2\right\} \end{split}$$

▶ Conditional Posteriors:

$$p(\mu \mid \sigma^2, \mathbf{y}_{1:n}) = \text{Normal}(\mu_n, \sigma^2/\kappa_n)$$

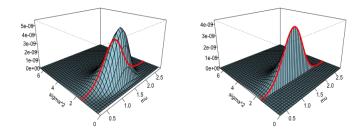
$$p(\sigma^2 \mid \mu, \mathbf{y}_{1:n}) = \text{Inv-Ga}[(\nu_n + 1)/2, {\kappa_n(\mu - \mu_n)^2 + \nu_n \sigma_n^2}/2]$$

4/18 8/71

(WB)

► Conditional Posterior of μ for given values of σ^2 :

$$\triangleright p(\mu \mid \sigma^2, \mathbf{y}_{1:n}) = \text{Normal}(\mu_n, \sigma^2/\kappa_n)$$



The blue surface shows the joint NIG posterior $p(\mu, \sigma^2 \mid \mathbf{y}_{1:n})$. The sliced red curve shows the conditional Normal posterior $p(\mu \mid \sigma^2, \mathbf{y}_{1:n})$ for $\sigma^2 \approx 1.7$.

5/18 9/

Normal Model Marginal Posteriors under Conjugate Prior

$$y_1, \ldots, y_n \stackrel{iid}{\sim} \text{Normal}(\mu, \sigma^2)$$
 with μ, σ^2 both unknown

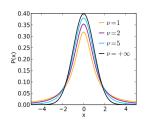
- Normal Likelihood: $p(\mathbf{y}_{1:n} \mid \mu, \sigma^2) \propto (\sigma^2)^{-\frac{n}{2}} \exp \left[-\frac{1}{2\sigma^2} \left\{ (n-1)s^2 + n(\overline{y} \mu)^2 \right\} \right]$
- ▶ Normal-Inverse-Gamma Prior: $(\mu, \sigma^2) \sim \text{NIG}\left(\mu_0, \sigma_0^2 / \kappa_0, \nu_0, \sigma_0^2\right)$

$$\begin{split} p(\mu,\sigma^2) &= p(\sigma^2) p(\mu \mid \sigma^2) = \text{Inv-Ga}\left(\sigma^2 \mid \nu_0/2, \nu_0 \sigma_0^2/2\right) \cdot \text{Normal}(\mu \mid \mu_0, \sigma^2/\kappa_0) \\ &\propto \left(\sigma^2\right)^{-\left(\frac{\nu_0}{2} + 1 + \frac{1}{2}\right)} \exp\left[-\frac{1}{2\sigma^2} \left\{\nu_0 \sigma_0^2 + \kappa_0 (\mu - \mu_0)^2\right\}\right] \end{split}$$

► Normal-Inverse-Gamma Posterior:

$$\begin{split} & p(\mu,\sigma^2\mid\mathbf{y}_{1:n})\\ & \propto \left(\sigma^2\right)^{-\left\{\frac{(\nu_0+n)}{2}+1+\frac{1}{2}\right\}} \exp\left[-\frac{1}{2\sigma^2}\left\{\nu_0\sigma_0^2+(n-1)s^2+\frac{n\kappa_0}{(n+\kappa_0)}(\overline{y}-\mu_0)^2+(\kappa_0+n)(\mu-\mu_n)^2\right\}\right]\\ & \equiv \mathrm{NIG}\left(\mu_n,\sigma_n^2/\kappa_n,\nu_n,\sigma_n^2\right), \quad \nu_n=(\nu_0+n), \quad \kappa_n=(\kappa_0+n), \quad \mu_n=(\kappa_0\mu_0+n\overline{y})/(\kappa_0+n),\\ & \sigma_n^2=\frac{1}{\nu_n}\left\{\nu_0\sigma_0^2+(n-1)s^2+\frac{n\kappa_0}{(n+\kappa_0)}(\overline{y}-\mu_0)^2\right\} \end{split}$$

► Marginal Posteriors:



► Student's t-distribution:

$$\begin{split} p(y\mid\nu) &= \frac{1}{\sqrt{\nu}\,\operatorname{Beta}\left(\frac{\nu+1}{2},\frac{1}{2}\right)} \cdot \left(1 + \frac{y^2}{\nu}\right)^{-\frac{(\nu+1)}{2}} \\ \mathbb{E}(y) &= 0, \quad \nu > 1, \quad \operatorname{var}(y) = \frac{\nu}{(\nu-2)}, \quad \nu > 2 \end{split}$$

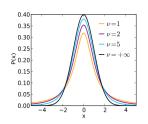
 $ightharpoonup p(y \mid \nu) \to \text{Normal}(0,1) \text{ as } \nu \to \infty.$

Scaled and Shifted t-distribution:

$$\begin{split} p(y\mid\nu,\mu,\sigma^2) &= \frac{1}{\sqrt{\nu}\,\operatorname{Beta}\left(\frac{\nu+1}{2},\frac{1}{2}\right)\,\,\sigma}.\left\{1 + \frac{1}{\nu}\left(\frac{y-\mu}{\sigma}\right)^2\right\}^{-\frac{(\nu+1)}{2}}\\ \mathbb{E}(y) &= \mu, \quad \nu > 1, \qquad \operatorname{var}(y) = \sigma^2\frac{\nu}{(\nu-2)}, \quad \nu > 2 \end{split}$$

▶ $p(y \mid \nu, 0, 1) \equiv$ Student's t-distribution.

7/18 11/71



▶ Student's t-distribution:

$$\begin{split} p(y\mid\nu) &= \frac{1}{\sqrt{\nu}\,\operatorname{Beta}\left(\frac{\nu+1}{2},\frac{1}{2}\right)}.\left(1+\frac{y^2}{\nu}\right)^{-\frac{(\nu+1)}{2}}\\ \mathbb{E}(y) &= 0, \quad \nu > 1, \quad \operatorname{var}(y) = \frac{\nu}{(\nu-2)}, \quad \nu > 2 \end{split}$$

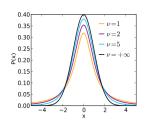
$$\blacktriangleright \ p(y\mid\nu) \to \operatorname{Normal}(0,1) \text{ as } \nu \to \infty. \end{split}$$

Scaled and Shifted t-distribution:

$$\begin{split} p(y\mid\nu,\mu,\sigma^2) &= \frac{1}{\sqrt{\nu}\,\operatorname{Beta}\left(\frac{\nu+1}{2},\frac{1}{2}\right)\,\sigma}.\left\{1 + \frac{1}{\nu}\left(\frac{y-\mu}{\sigma}\right)^2\right\}^{-\frac{(\nu+1)}{2}}\\ \mathbb{E}(y) &= \mu, \quad \nu > 1, \qquad \operatorname{var}(y) = \sigma^2\frac{\nu}{(\nu-2)}, \quad \nu > 2 \end{split}$$

▶ $p(y \mid \nu, 0, 1) \equiv$ Student's t-distribution.

7/18 12/71



► Student's t-distribution:

$$\begin{split} p(y\mid\nu) &= \frac{1}{\sqrt{\nu}\;\mathrm{Beta}\left(\frac{\nu+1}{2},\frac{1}{2}\right)}.\left(1+\frac{y^2}{\nu}\right)^{-\frac{(\nu+1)}{2}}\\ \mathbb{E}(y) &= 0, \quad \nu > 1, \quad \mathrm{var}(y) = \frac{\nu}{(\nu-2)}, \quad \nu > 2 \end{split}$$

$$\blacktriangleright \; p(y\mid\nu) \to \mathrm{Normal}(0,1) \; \mathrm{as}\; \nu \to \infty. \end{split}$$

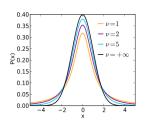
► Scaled and Shifted t-distribution:

$$\begin{split} p(y\mid\nu,\mu,\sigma^2) &= \frac{1}{\sqrt{\nu}\;\mathrm{Beta}\left(\frac{\nu+1}{2},\frac{1}{2}\right)\;\sigma}.\left\{1+\frac{1}{\nu}\left(\frac{y-\mu}{\sigma}\right)^2\right\}^{-\frac{(\nu+1)}{2}}\\ \mathbb{E}(y) &= \mu, \quad \nu > 1, \quad \mathrm{var}(y) = \sigma^2\frac{\nu}{(\nu-2)}, \quad \nu > 2 \end{split}$$

 $p(y \mid \nu, 0, 1) \equiv$ Student's t-distribution.

7/18 13/71

Normal Model t-distributions



Student's t-distribution:

$$\begin{split} p(y\mid\nu) &= \frac{1}{\sqrt{\nu}\operatorname{Beta}\left(\frac{\nu+1}{2},\frac{1}{2}\right)} \cdot \left(1 + \frac{y^2}{\nu}\right)^{-\frac{(\nu+1)}{2}} \\ \mathbb{E}(y) &= 0, \quad \nu > 1, \quad \operatorname{var}(y) = \frac{\nu}{(\nu-2)}, \quad \nu > 2 \end{split}$$

▶ $p(y \mid \nu) \rightarrow \text{Normal}(0, 1) \text{ as } \nu \rightarrow \infty.$

► Scaled and Shifted t-distribution:

$$\begin{split} p(y\mid\nu,\mu,\sigma^2) &= \frac{1}{\sqrt{\nu}\;\mathrm{Beta}\left(\frac{\nu+1}{2},\frac{1}{2}\right)\;\sigma}.\left\{1+\frac{1}{\nu}\left(\frac{y-\mu}{\sigma}\right)^2\right\}^{-\frac{(\nu+1)}{2}}\\ \mathbb{E}(y) &= \mu, \quad \nu > 1, \qquad \mathrm{var}(y) = \sigma^2\frac{\nu}{(\nu-2)}, \quad \nu > 2 \end{split}$$

▶ $p(y \mid \nu, 0, 1) \equiv$ Student's t-distribution.

Normal Model Predictive Distribution under Conjugate Prior

$$y_1, \ldots, y_n \overset{iid}{\sim} \operatorname{Normal}(\mu, \sigma^2)$$
 with μ, σ^2 both unknown

- Normal Likelihood: $p(\mathbf{y}_{1:n} \mid \mu, \sigma^2) \propto (\sigma^2)^{-\frac{n}{2}} \exp \left[-\frac{1}{2\sigma^2} \left\{ (n-1)s^2 + n(\overline{y} \mu)^2 \right\} \right]$
- ▶ Normal-Inverse-Gamma Prior: $(\mu, \sigma^2) \sim \text{NIG}(\mu_0, \sigma_0^2 / \kappa_0, \nu_0, \sigma_0^2)$

$$\begin{split} p(\mu,\sigma^2) &= p(\sigma^2) p(\mu \mid \sigma^2) = \text{Inv-Ga}\left(\sigma^2 \mid \nu_0/2,\nu_0\sigma_0^2/2\right) \cdot \text{Normal}(\mu \mid \mu_0,\sigma^2/\kappa_0) \\ &\propto \left(\sigma^2\right)^{-\left(\frac{\nu_0}{2}+1+\frac{1}{2}\right)} \exp\left[-\frac{1}{2\sigma^2}\left\{\nu_0\sigma_0^2 + \kappa_0(\mu-\mu_0)^2\right\}\right] \end{split}$$

▶ Normal-Inverse-Gamma Posterior:

$$\begin{split} &p(\mu,\sigma^2\mid\mathbf{y}_{1:n})\\ &\propto \left(\sigma^2\right)^{-\left\{\frac{(\nu_0+n)}{2}+1+\frac{1}{2}\right\}} \exp\left[-\frac{1}{2\sigma^2}\left\{\nu_0\sigma_0^2+(n-1)s^2+\frac{n\kappa_0}{(n+\kappa_0)}(\overline{y}-\mu_0)^2+(\kappa_0+n)(\mu-\mu_n)^2\right\}\right]\\ &\equiv \mathrm{NIG}\left(\mu_n,\sigma_n^2/\kappa_n,\nu_n,\sigma_n^2\right),\quad \nu_n=(\nu_0+n),\quad \kappa_n=(\kappa_0+n),\quad \mu_n=(\kappa_0\mu_0+n\overline{y})/(\kappa_0+n),\\ &\sigma_n^2=\frac{1}{\nu_n}\left\{\nu_0\sigma_0^2+(n-1)s^2+\frac{n\kappa_0}{(n+\kappa_0)}(\overline{y}-\mu_0)^2\right\} \end{split}$$

▶ Predictive Distribution:

8/18 15/71

$$y_1, \ldots, y_n \stackrel{iid}{\sim} \text{Normal}(\mu, \sigma^2)$$
 with μ, σ^2 both unknown

- ▶ Normal Likelihood: $p(\mathbf{y}_{1:n} \mid \mu, \sigma^2) \propto (\sigma^2)^{-\frac{n}{2}} \exp \left[-\frac{1}{2\sigma^2} \left\{ (n-1)s^2 + n(\overline{y} \mu)^2 \right\} \right]$
- ▶ Non-informative Improper Prior: $p(\mu, \sigma^2) \propto 1 \cdot \frac{1}{\sigma^2}$
- ▶ Posterior: $p(\mu, \sigma^2 \mid \mathbf{y}_{1:n}) \propto (\sigma^2)^{-(\frac{n}{2}+1)} \exp \left[-\frac{1}{2\sigma^2} \left\{ (n-1)s^2 + n(\mu \overline{y})^2 \right\} \right]$
- ightharpoonup A-posteriori μ and σ^2 are dependent but uncorrelated.

$$y_1, \ldots, y_n \stackrel{iid}{\sim} \text{Normal}(\mu, \sigma^2)$$
 with μ, σ^2 both unknown

- Normal Likelihood: $p(\mathbf{y}_{1:n} \mid \mu, \sigma^2) \propto (\sigma^2)^{-\frac{n}{2}} \exp \left[-\frac{1}{2\sigma^2} \left\{ (n-1)s^2 + n(\overline{y} \mu)^2 \right\} \right]$
- Non-informative Improper Prior: $p(\mu, \sigma^2) \propto 1 \cdot \frac{1}{\sigma^2}$
- ▶ Posterior: $p(\mu, \sigma^2 \mid \mathbf{y}_{1:n}) \propto (\sigma^2)^{-(\frac{n}{2}+1)} \exp\left[-\frac{1}{2\sigma^2}\left\{(n-1)s^2 + n(\mu \overline{y})^2\right\}\right]$

ightharpoonup A-posteriori μ and σ^2 are dependent but uncorrelated.

$$y_1, \ldots, y_n \stackrel{iid}{\sim} \text{Normal}(\mu, \sigma^2)$$
 with μ, σ^2 both unknown

- ▶ Normal Likelihood: $p(\mathbf{y}_{1:n} \mid \mu, \sigma^2) \propto (\sigma^2)^{-\frac{n}{2}} \exp \left[-\frac{1}{2\sigma^2} \left\{ (n-1)s^2 + n(\overline{y} \mu)^2 \right\} \right]$
- Non-informative Improper Prior: $p(\mu, \sigma^2) \propto 1 \cdot \frac{1}{\sigma^2}$
- ▶ Posterior: $p(\mu, \sigma^2 \mid \mathbf{y}_{1:n}) \propto (\sigma^2)^{-(\frac{n}{2}+1)} \exp\left[-\frac{1}{2\sigma^2}\left\{(n-1)s^2 + n(\mu-\overline{y})^2\right\}\right]$
- A-posteriori μ and σ^2 are dependent but uncorrelated.

$$y_1, \ldots, y_n \stackrel{iid}{\sim} \text{Normal}(\mu, \sigma^2)$$
 with μ, σ^2 both unknown

- ▶ Normal Likelihood: $p(\mathbf{y}_{1:n} \mid \mu, \sigma^2) \propto (\sigma^2)^{-\frac{n}{2}} \exp \left[-\frac{1}{2\sigma^2} \left\{ (n-1)s^2 + n(\overline{y} \mu)^2 \right\} \right]$
- Non-informative Improper Prior: $p(\mu, \sigma^2) \propto 1 \cdot \frac{1}{\sigma^2}$
- ▶ Posterior: $p(\mu, \sigma^2 \mid \mathbf{y}_{1:n}) \propto \left(\sigma^2\right)^{-\left(\frac{n}{2}+1\right)} \exp\left[-\frac{1}{2\sigma^2}\left\{(n-1)s^2 + n(\mu \overline{y})^2\right\}\right]$
- ▶ A-posteriori μ and σ^2 are dependent but uncorrelated. **HW.4.1b**

9/18 19/71

$$y_1, \ldots, y_n \stackrel{iid}{\sim} \text{Normal}(\mu, \sigma^2)$$
 with μ, σ^2 both unknown

- Normal Likelihood: $p(\mathbf{y}_{1:n} \mid \mu, \sigma^2) \propto (\sigma^2)^{-\frac{n}{2}} \exp \left[-\frac{1}{2\sigma^2} \left\{ (n-1)s^2 + n(\overline{y} \mu)^2 \right\} \right]$
- Non-informative Improper Prior: $p(\mu, \sigma^2) \propto 1 \cdot \frac{1}{\sigma^2}$
- ▶ Posterior: $p(\mu, \sigma^2 \mid \mathbf{y}_{1:n}) \propto (\sigma^2)^{-(\frac{n}{2}+1)} \exp\left[-\frac{1}{2\sigma^2}\left\{(n-1)s^2 + n(\mu \overline{y})^2\right\}\right]$
- A-posteriori μ and σ^2 are dependent but uncorrelated.
- **▶** Conditional Posteriors:
 - $ightharpoonup p(\mu \mid \sigma^2, \mathbf{y}_{1:n}) = \operatorname{Normal}(\overline{y}, \sigma^2/n)$
- **▶** Marginal Posteriors

▶ Predictive Distribution: $p(y_{new} | \mathbf{y}_{1:n}) = ?$

9/18 20/71

$$y_1, \ldots, y_n \stackrel{iid}{\sim} \text{Normal}(\mu, \sigma^2)$$
 with μ, σ^2 both unknown

- Normal Likelihood: $p(\mathbf{y}_{1:n} \mid \mu, \sigma^2) \propto (\sigma^2)^{-\frac{n}{2}} \exp \left[-\frac{1}{2\sigma^2} \left\{ (n-1)s^2 + n(\overline{y} \mu)^2 \right\} \right]$
- ▶ Non-informative Improper Prior: $p(\mu, \sigma^2) \propto 1 \cdot \frac{1}{\sigma^2}$
- ▶ Posterior: $p(\mu, \sigma^2 \mid \mathbf{y}_{1:n}) \propto (\sigma^2)^{-(\frac{n}{2}+1)} \exp\left[-\frac{1}{2\sigma^2}\left\{(n-1)s^2 + n(\mu \overline{y})^2\right\}\right]$
- ▶ A-posteriori μ and σ^2 are dependent but uncorrelated.
- **▶** Conditional Posteriors:
 - $\blacktriangleright p(\mu \mid \sigma^2, \mathbf{y}_{1:n}) = \text{Normal}(\overline{y}, \sigma^2/n)$
- **▶** Marginal Posteriors:

▶ Predictive Distribution: $p(y_{new} \mid \mathbf{v}_{1:n}) = ?$

9/18 21/71

$$y_1, \ldots, y_n \stackrel{iid}{\sim} \text{Normal}(\mu, \sigma^2)$$
 with μ, σ^2 both unknown

- Normal Likelihood: $p(\mathbf{y}_{1:n} \mid \mu, \sigma^2) \propto (\sigma^2)^{-\frac{n}{2}} \exp \left[-\frac{1}{2\sigma^2} \left\{ (n-1)s^2 + n(\overline{y} \mu)^2 \right\} \right]$
- ▶ Non-informative Improper Prior: $p(\mu, \sigma^2) \propto 1 \cdot \frac{1}{\sigma^2}$
- ▶ **Posterior:** $p(\mu, \sigma^2 \mid \mathbf{y}_{1:n}) \propto (\sigma^2)^{-(\frac{n}{2}+1)} \exp \left[-\frac{1}{2\sigma^2} \left\{ (n-1)s^2 + n(\mu \overline{y})^2 \right\} \right]$
- ▶ A-posteriori μ and σ^2 are dependent but uncorrelated.
- **▶** Conditional Posteriors:
 - $ightharpoonup p(\mu \mid \sigma^2, \mathbf{y}_{1:n}) = \text{Normal}(\overline{y}, \sigma^2/n)$
- Marginal Posteriors:
 - $\blacktriangleright p(\mu \mid \mathbf{y}_{1:n}) = t_{n-1}\left(\overline{y}, \frac{s^2}{n}\right) \quad \blacktriangleright p(\sigma^2 \mid \mathbf{y}_{1:n}) = \text{Inv-Ga}\left\{\frac{(n-1)}{2}, \frac{(n-1)s^2}{2}\right\}$

▶ Predictive Distribution: $p(y_{new} | \mathbf{y}_{1:n}) = ?$

9/18 22/71

$$y_1, \ldots, y_n \stackrel{iid}{\sim} \text{Normal}(\mu, \sigma^2)$$
 with μ, σ^2 both unknown

- Normal Likelihood: $p(\mathbf{y}_{1:n} \mid \mu, \sigma^2) \propto (\sigma^2)^{-\frac{n}{2}} \exp \left[-\frac{1}{2\sigma^2} \left\{ (n-1)s^2 + n(\overline{y} \mu)^2 \right\} \right]$
- ▶ Non-informative Improper Prior: $p(\mu, \sigma^2) \propto 1 \cdot \frac{1}{\sigma^2}$
- ▶ **Posterior:** $p(\mu, \sigma^2 \mid \mathbf{y}_{1:n}) \propto (\sigma^2)^{-(\frac{n}{2}+1)} \exp \left[-\frac{1}{2\sigma^2} \left\{ (n-1)s^2 + n(\mu \overline{y})^2 \right\} \right]$
- ▶ A-posteriori μ and σ^2 are dependent but uncorrelated.
- **▶** Conditional Posteriors:
 - $\blacktriangleright p(\mu \mid \sigma^2, \mathbf{y}_{1:n}) = \text{Normal}(\overline{y}, \sigma^2/n)$
- **▶** Marginal Posteriors:
 - $\blacktriangleright p(\mu \mid \mathbf{y}_{1:n}) = t_{n-1}\left(\overline{y}, \frac{s^2}{n}\right) \quad \blacktriangleright p(\sigma^2 \mid \mathbf{y}_{1:n}) = \text{Inv-Ga}\left\{\frac{(n-1)}{2}, \frac{(n-1)s^2}{2}\right\}$

▶ Predictive Distribution: $p(y_{new} | \mathbf{y}_{1:n}) = ?$

9/18 23/71

$$y_1, \ldots, y_n \stackrel{iid}{\sim} \text{Normal}(\mu, \sigma^2)$$
 with μ, σ^2 both unknown

- Normal Likelihood: $p(\mathbf{y}_{1:n} \mid \mu, \sigma^2) \propto (\sigma^2)^{-\frac{n}{2}} \exp \left[-\frac{1}{2\sigma^2} \left\{ (n-1)s^2 + n(\overline{y} \mu)^2 \right\} \right]$
- $\blacktriangleright \ \ Independent \ Normal \times Inverse\text{-}Gamma \ Prior:$

$$(\mu, \sigma^2) \sim ext{Normal}(\mu_0, \sigma_0^2) ext{ Inv-Ga}(a_0, b_0)$$

$$\begin{split} p(\mu,\sigma^2) &= p(\mu)p(\sigma^2) = \text{Normal}(\mu \mid \mu_0,\sigma_0^2) \cdot \text{Inv-Ga}\left(\sigma^2 \mid a_0,b_0\right) \\ &\propto \exp\left\{-\frac{1}{2\sigma_0^2}(\mu-\mu_0)^2\right\}\left(\sigma^2\right)^{-(a_0+1)} \exp\left(-\frac{b_0}{\sigma^2}\right) \end{split}$$

Posterior:

$$p(\mu, \sigma^2 \mid \mathbf{y}_{1:n}) \propto \left(\sigma^2\right)^{-\left(a_0 + \frac{n}{2} + 1\right)} \exp\left[-\frac{1}{\sigma^2} \left\{b_0 + \frac{(n-1)s^2}{2} + \frac{n(\overline{y} - \mu)^2}{2}\right\}\right] \exp\left\{-\frac{1}{2\sigma_0^2} (\mu - \mu_0)^2\right\}$$

10/18 24/71

$$y_1, \ldots, y_n \stackrel{iid}{\sim} \text{Normal}(\mu, \sigma^2)$$
 with μ, σ^2 both unknown

- Normal Likelihood: $p(\mathbf{y}_{1:n} \mid \mu, \sigma^2) \propto (\sigma^2)^{-\frac{n}{2}} \exp \left[-\frac{1}{2\sigma^2} \left\{ (n-1)s^2 + n(\overline{y} \mu)^2 \right\} \right]$
- ▶ Independent Normal × Inverse-Gamma Prior:

$$\begin{split} & \left(\mu,\sigma^2\right) \sim \text{Normal}\big(\mu_0,\sigma_0^2\big) \text{ Inv-Ga}\big(a_0,b_0\big) \\ & p(\mu,\sigma^2) = p(\mu)p(\sigma^2) = \text{Normal}\big(\mu \mid \mu_0,\sigma_0^2\big) \cdot \text{Inv-Ga}\left(\sigma^2 \mid a_0,b_0\right) \\ & \propto \exp\left\{-\frac{1}{2\sigma_0^2}(\mu-\mu_0)^2\right\}\left(\sigma^2\right)^{-(a_0+1)} \exp\left(-\frac{b_0}{\sigma^2}\right) \end{split}$$

Posterior:

$$p(\mu, \sigma^2 \mid \mathbf{y}_{1:n}) \propto \left(\sigma^2\right)^{-\left(a_0 + \frac{n}{2} + 1\right)} \exp\left[-\frac{1}{\sigma^2} \left\{b_0 + \frac{(n-1)s^2}{2} + \frac{n(\overline{y} - \mu)^2}{2}\right\}\right] \exp\left\{-\frac{1}{2\sigma_0^2} (\mu - \mu_0)^2\right\}$$

10/18 25/71

$$y_1, \ldots, y_n \stackrel{iid}{\sim} \text{Normal}(\mu, \sigma^2)$$
 with μ, σ^2 both unknown

- Normal Likelihood: $p(\mathbf{y}_{1:n} \mid \mu, \sigma^2) \propto (\sigma^2)^{-\frac{n}{2}} \exp \left[-\frac{1}{2\sigma^2} \left\{ (n-1)s^2 + n(\overline{y} \mu)^2 \right\} \right]$
- ightharpoonup Independent Normal imes Inverse-Gamma Prior:

$$\begin{split} & \left(\mu,\sigma^2\right) \sim \text{Normal}\big(\mu_0,\sigma_0^2\big) \text{ Inv-Ga}\big(a_0,b_0\big) \\ & p(\mu,\sigma^2) = p(\mu)p(\sigma^2) = \text{Normal}\big(\mu \mid \mu_0,\sigma_0^2\big) \cdot \text{Inv-Ga}\left(\sigma^2 \mid a_0,b_0\right) \\ & \propto \exp\left\{-\frac{1}{2\sigma_0^2}(\mu-\mu_0)^2\right\}\left(\sigma^2\right)^{-(a_0+1)} \exp\left(-\frac{b_0}{\sigma^2}\right) \end{split}$$

▶ Posterior:

$$p(\mu,\sigma^2\mid\mathbf{y}_{1:n})\propto \left(\sigma^2\right)^{-\left(a_0+\frac{n}{2}+1\right)}\exp\left[-\frac{1}{\sigma^2}\left\{b_0+\frac{(n-1)s^2}{2}+\frac{n(\overline{y}-\mu)^2}{2}\right\}\right]\exp\left\{-\frac{1}{2\sigma_0^2}(\mu-\mu_0)^2\right\}$$

10/18 26/71

$$y_1, \ldots, y_n \overset{iid}{\sim} \text{Normal}(\mu, \sigma^2)$$
 with μ, σ^2 both unknown

- Normal Likelihood: $p(\mathbf{y}_{1:n} \mid \mu, \sigma^2) \propto (\sigma^2)^{-\frac{n}{2}} \exp \left[-\frac{1}{2\sigma^2} \left\{ (n-1)s^2 + n(\overline{y} \mu)^2 \right\} \right]$
- ightharpoonup Independent Normal imes Inverse-Gamma Prior:

$$\begin{split} &(\mu,\sigma^2) \sim \text{Normal}\big(\mu_0,\sigma_0^2\big) \text{ Inv-Ga}\big(a_0,b_0\big) \\ &p(\mu,\sigma^2) = p(\mu)p(\sigma^2) = \text{Normal}\big(\mu \mid \mu_0,\sigma_0^2\big) \cdot \text{Inv-Ga}\left(\sigma^2 \mid a_0,b_0\right) \\ &\propto \exp\left\{-\frac{1}{2\sigma_0^2}(\mu-\mu_0)^2\right\}\left(\sigma^2\right)^{-(a_0+1)} \exp\left(-\frac{b_0}{\sigma^2}\right) \end{split}$$

▶ Posterior:

$$p(\mu, \sigma^2 \mid \mathbf{y}_{1:n}) \propto \left(\sigma^2\right)^{-\left(a_0 + \frac{n}{2} + 1\right)} \exp\left[-\frac{1}{\sigma^2} \left\{b_0 + \frac{(n-1)s^2}{2} + \frac{n(\overline{y} - \mu)^2}{2}\right\}\right] \exp\left\{-\frac{1}{2\sigma_0^2} (\mu - \mu_0)^2\right\}$$

▶ A-priori μ and σ^2 are independent but a-posteriori they are not.

10/18 27/71

$$y_1, \ldots, y_n \stackrel{iid}{\sim} \text{Normal}(\mu, \sigma^2)$$
 with μ, σ^2 both unknown

- Normal Likelihood: $p(\mathbf{y}_{1:n} \mid \mu, \sigma^2) \propto (\sigma^2)^{-\frac{n}{2}} \exp \left[-\frac{1}{2\sigma^2} \left\{ (n-1)s^2 + n(\overline{y} \mu)^2 \right\} \right]$
- ightharpoonup Independent Normal imes Inverse-Gamma Prior:

$$\begin{split} & (\mu, \sigma^2) \sim \text{Normal}(\mu_0, \sigma_0^2) \text{ Inv-Ga}(a_0, b_0) \\ & p(\mu, \sigma^2) = p(\mu) p(\sigma^2) = \text{Normal}(\mu \mid \mu_0, \sigma_0^2) \cdot \text{Inv-Ga}\left(\sigma^2 \mid a_0, b_0\right) \\ & \propto \exp\left\{-\frac{1}{2\sigma_0^2}(\mu - \mu_0)^2\right\} \left(\sigma^2\right)^{-(a_0 + 1)} \exp\left(-\frac{b_0}{\sigma^2}\right) \end{split}$$

▶ Posterior:

$$p(\mu, \sigma^2 \mid \mathbf{y}_{1:n}) \propto \left(\sigma^2\right)^{-\left(a_0 + \frac{n}{2} + 1\right)} \exp\left[-\frac{1}{\sigma^2} \left\{b_0 + \frac{(n-1)s^2}{2} + \frac{n(\overline{y} - \mu)^2}{2}\right\}\right] \exp\left\{-\frac{1}{2\sigma_0^2} (\mu - \mu_0)^2\right\}$$

▶ Conditional Posteriors:

$$\blacktriangleright p(\mu \mid \sigma^2, \mathbf{y}_{1:n}) = \text{Normal}(\mu_n, \sigma_n^2), \quad \mu_n = \sigma_n^2 \left(\frac{\mu_0}{\sigma_0^2} + \frac{n\overline{y}}{\sigma^2}\right), \ \sigma_n^2 = \left(\frac{1}{\sigma_0^2} + \frac{n}{\sigma^2}\right)^{-1},$$

10/18 28/71

$$\mathbf{y}_1, \dots, \mathbf{y}_n \overset{iid}{\sim} \mathsf{MVN}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$
 with $\boldsymbol{\mu}^{d \times 1}, \boldsymbol{\Sigma}^{d \times d}$ both unknown

- ▶ MVN Likelihood: $p(\mathbf{y}_{1:n} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(\mathbf{y}_{i} \boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{y}_{i} \boldsymbol{\mu})\right\}$ $\propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left[-\frac{1}{2}\left\{n(\boldsymbol{\mu} \overline{\mathbf{y}})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} \overline{\mathbf{y}}) + \sum_{i=1}^{n}(\mathbf{y}_{i} \overline{\mathbf{y}})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{y}_{i} \overline{\mathbf{y}})\right\}\right]$ $\propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left(-\frac{1}{2}\left[n(\boldsymbol{\mu} \overline{\mathbf{y}})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} \overline{\mathbf{y}}) + \operatorname{trace}\{\boldsymbol{\Sigma}^{-1}\sum_{i=1}^{n}(\mathbf{y}_{i} \overline{\mathbf{y}})(\mathbf{y}_{i} \overline{\mathbf{y}})^{\mathrm{T}}\}\right]\right)$ $\propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left[-\frac{1}{2}\left\{n(\boldsymbol{\mu} \overline{\mathbf{y}})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} \overline{\mathbf{y}}) + \operatorname{trace}(\boldsymbol{\Sigma}^{-1}\mathbf{S})\right\}\right]$
- ▶ Wishart Distribution: $\mathbf{W} \sim W_d(n, \Omega)$ has the pdf $p(\mathbf{W}) \propto |\mathbf{W}|^{\frac{n-d-1}{2}} \exp\left\{-\frac{1}{2}\mathrm{trace}(\mathbf{W}\Omega^{-1})\right\}, \quad n > d-1, \; \Omega > 0, \; \mathbf{W} \in \mathbf{M}^+.$

▶ Inverse-Wishart Distribution: If $W \sim W_d(n, \Omega)$ and $V = W^{-1}$ and $\Lambda = \Omega^{-1}$, then $V \sim IW_d(n, \Lambda)$ and has the pdf

$$\mathbf{y}_1, \dots, \mathbf{y}_n \overset{iid}{\sim} \text{MVN}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$
 with $\boldsymbol{\mu}^{d \times 1}, \boldsymbol{\Sigma}^{d \times d}$ both unknown

- ▶ MVN Likelihood: $p(\mathbf{y}_{1:n} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(\mathbf{y}_{i} \boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{y}_{i} \boldsymbol{\mu})\right\}$ $\propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left[-\frac{1}{2}\left\{n(\boldsymbol{\mu} - \overline{\mathbf{y}})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \overline{\mathbf{y}}) + \operatorname{trace}(\boldsymbol{\Sigma}^{-1}\mathbf{S})\right\}\right]$
- ▶ Wishart Distribution: $\mathbf{W} \sim W_d(n, \mathbf{\Omega})$ has the pdf $p(\mathbf{W}) \propto |\mathbf{W}|^{\frac{n-d-1}{2}} \exp\left\{-\frac{1}{2}\mathrm{trace}(\mathbf{W}\mathbf{\Omega}^{-1})\right\}, \quad n > d-1, \; \mathbf{\Omega} > 0, \; \mathbf{W} \in \mathbf{M}^+.$ \mathbf{M}^+ is the set of all positive definite matrices.
 - ▶ If $z_1, \ldots, z_n \sim \text{Normal}(0, \sigma^2)$, then $\sum_{i=1}^n z_i^2 \sim \sigma^2 \chi_n^2$
 - ightharpoonup If $\mathbf{z}_1, \dots, \mathbf{z}_n \sim \text{MVN}_d(\mathbf{0}, \Omega)$, then $\mathbf{Z}^T \mathbf{Z} \sim W_d(n, \Omega)$.
 - ▶ If $\mathbf{W} \sim W_d(n, \mathbf{\Omega})$, then $\mathbb{E}(\mathbf{W}) = n\mathbf{\Omega}$.
 - ightharpoonup If $\mathbf{W} \sim W_d(n, \mathbf{\Omega})$, then, for any $\mathbf{B}^{q \times d}$, $\mathbf{B}\mathbf{W}\mathbf{B}^{\mathrm{T}} \sim W_d(n, \mathbf{B}\mathbf{\Omega}\mathbf{B}^{\mathrm{T}})$.
 - ▶ If $\mathbf{W} = ((w_{ij})) \sim W_d(n, \mathbf{\Omega})$ with $\mathbf{\Omega} = ((\omega_{ij}))$, then $w_{ii} \sim \omega_{ii} \chi_n^2$.
 - ▶ If $\mathbf{W}_j \stackrel{ina}{\sim} W_d(n_j, \Omega)$, then $\sum_{j=1}^J \mathbf{W}_j \sim W_d(\sum_{j=1}^J n_j, \Omega)$.
- ▶ Inverse-Wishart Distribution: If $W \sim W_d(n, \Omega)$ and $V = W^{-1}$ and $\Lambda = \Omega^{-1}$, then $V \sim IW_d(n, \Lambda)$ and has the pdf

 $p(\mathbf{V}) \propto |\mathbf{V}|^{-\frac{n+d+1}{2}} \exp\left\{-\frac{1}{2}\operatorname{trace}(\mathbf{V}^{-1}\mathbf{\Lambda})\right\}, \quad n > d-1, \ \mathbf{\Lambda} > 0, \ \mathbf{V} \in \mathbf{M}^+.$

11/18 30/71

$$\mathbf{y}_1, \dots, \mathbf{y}_n \overset{iid}{\sim} \text{MVN}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$
 with $\boldsymbol{\mu}^{d \times 1}, \boldsymbol{\Sigma}^{d \times d}$ both unknown

- ▶ MVN Likelihood: $p(\mathbf{y}_{1:n} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(\mathbf{y}_{i} \boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{y}_{i} \boldsymbol{\mu})\right\}$ $\propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left[-\frac{1}{2}\left\{n(\boldsymbol{\mu} - \overline{\mathbf{y}})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \overline{\mathbf{y}}) + \operatorname{trace}(\boldsymbol{\Sigma}^{-1}\mathbf{S})\right\}\right]$
- ▶ Wishart Distribution: $\mathbf{W} \sim W_d(n, \mathbf{\Omega})$ has the pdf $p(\mathbf{W}) \propto |\mathbf{W}|^{\frac{n-d-1}{2}} \exp\left\{-\frac{1}{2}\mathrm{trace}(\mathbf{W}\mathbf{\Omega}^{-1})\right\}, \quad n > d-1, \; \mathbf{\Omega} > 0, \; \mathbf{W} \in \mathbf{M}^+.$
 - ▶ If $z_1, \ldots, z_n \sim \text{Normal}(0, \sigma^2)$, then $\sum_{i=1}^n z_i^2 \sim \sigma^2 \chi_n^2$.
 - ightharpoonup If $\mathbf{z}_1, \dots, \mathbf{z}_n \sim \text{MVN}_d(\mathbf{0}, \mathbf{\Omega})$, then $\mathbf{Z}^T \mathbf{Z} \sim W_d(n, \mathbf{\Omega})$.
 - ▶ If $\mathbf{W} \sim W_d(n, \mathbf{\Omega})$, then $\mathbb{E}(\mathbf{W}) = n\mathbf{\Omega}$.
 - ▶ If $\mathbf{W} \sim W_d(n, \mathbf{\Omega})$, then, for any $\mathbf{B}^{q \times d}$, $\mathbf{B}\mathbf{W}\mathbf{B}^{\mathrm{T}} \sim W_q(n, \mathbf{B}\mathbf{\Omega}\mathbf{B}^{\mathrm{T}})$.
 - ▶ If $\mathbf{W} = ((w_{ij})) \sim W_d(n, \mathbf{\Omega})$ with $\mathbf{\Omega} = ((\omega_{ij}))$, then $w_{ii} \sim \omega_{ii} \chi_n^2$.
 - ▶ If $\mathbf{W}_j \stackrel{ind}{\sim} W_d(n_j, \mathbf{\Omega})$, then $\sum_{j=1}^J \mathbf{W}_j \sim W_d(\sum_{j=1}^J n_j, \mathbf{\Omega})$
- ▶ Inverse-Wishart Distribution: If $\mathbf{W} \sim W_d(n, \mathbf{\Omega})$ and $\mathbf{V} = \mathbf{W}^{-1}$ and $\mathbf{\Lambda} = \mathbf{\Omega}^{-1}$, then $\mathbf{V} \sim \mathrm{IW}_d(n, \mathbf{\Lambda})$ and has the pdf
 - $p(\mathbf{V}) \propto |\mathbf{V}|^{-\frac{n+\alpha+1}{2}} \exp\left\{-\frac{1}{2}\operatorname{trace}(\mathbf{V}^{-1}\mathbf{\Lambda})\right\}, \quad n > d-1, \ \mathbf{\Lambda} > 0, \ \mathbf{V} \in \mathbf{M}^+.$

11/18 31/71

$$\mathbf{y}_1, \dots, \mathbf{y}_n \overset{iid}{\sim} \text{MVN}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$
 with $\boldsymbol{\mu}^{d \times 1}, \boldsymbol{\Sigma}^{d \times d}$ both unknown

- ▶ MVN Likelihood: $p(\mathbf{y}_{1:n} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(\mathbf{y}_{i} \boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{y}_{i} \boldsymbol{\mu})\right\}$ $\propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left[-\frac{1}{2}\left\{n(\boldsymbol{\mu} - \overline{\mathbf{y}})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \overline{\mathbf{y}}) + \operatorname{trace}(\boldsymbol{\Sigma}^{-1}\mathbf{S})\right\}\right]$
- ▶ Wishart Distribution: $W \sim W_d(n, \Omega)$ has the pdf

$$p(\mathbf{W}) \propto |\mathbf{W}|^{\frac{n-d-1}{2}} \exp\left\{-\frac{1}{2} \text{trace}(\mathbf{W} \mathbf{\Omega}^{-1})\right\}, \quad n>d-1, \; \mathbf{\Omega}>0, \; \mathbf{W} \in \mathbf{M}^+.$$

- ▶ If $z_1, \ldots, z_n \sim \text{Normal}(0, \sigma^2)$, then $\sum_{i=1}^n z_i^2 \sim \sigma^2 \chi_n^2$.
- ▶ If $\mathbf{z}_1, \dots, \mathbf{z}_n \sim \text{MVN}_d(\mathbf{0}, \mathbf{\Omega})$, then $\mathbf{Z}^T\mathbf{Z} \sim W_d(n, \mathbf{\Omega})$. Here

$$\mathbf{Z} = \begin{pmatrix} \mathbf{z}_{1}^{1} \\ \mathbf{z}_{2}^{T} \\ \vdots \\ \mathbf{z}_{n}^{T} \end{pmatrix} = \begin{bmatrix} z_{11} & z_{12} & \dots & z_{1d} \\ z_{21} & z_{22} & \dots & z_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n1} & z_{n2} & \dots & z_{nd} \end{bmatrix}^{n \times d}$$

$$\mathbf{Z}^{\mathrm{T}}\mathbf{Z} = \sum_{i=1}^{n} \mathbf{z}_{i}\mathbf{z}_{i}^{\mathrm{T}} = \begin{bmatrix} \sum_{i=1}^{n} z_{i1}^{2} & \sum_{i=1}^{n} z_{i1}z_{i2} & \cdots & \sum_{i=1}^{n} z_{i1}z_{id} \\ \sum_{i=1}^{n} z_{i1}z_{i2} & \sum_{i=1}^{n} z_{i2}^{2} & \cdots & \sum_{i=1}^{n} z_{i2}z_{id} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} z_{i1}z_{id} & \sum_{i=1}^{n} z_{i2}z_{id} & \cdots & \sum_{i=1}^{n} z_{id}^{2} \end{bmatrix}^{d \times d}$$

- ▶ If $\mathbf{W} \sim W_d(n, \mathbf{\Omega})$, then $\mathbb{E}(\mathbf{W}) = n\mathbf{\Omega}$.
- ▶ If $W \sim W_d(n, \Omega)$, then, for any $B^{q \wedge u}$, $BWB^+ \sim W_q(n, B\Omega B^+)$.

$$\mathbf{y}_1, \dots, \mathbf{y}_n \overset{iid}{\sim} \text{MVN}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$
 with $\boldsymbol{\mu}^{d \times 1}, \boldsymbol{\Sigma}^{d \times d}$ both unknown

- ▶ MVN Likelihood: $p(\mathbf{y}_{1:n} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(\mathbf{y}_{i} \boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{y}_{i} \boldsymbol{\mu})\right\}$ $\propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left[-\frac{1}{2}\left\{n(\boldsymbol{\mu} - \overline{\mathbf{y}})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \overline{\mathbf{y}}) + \operatorname{trace}(\boldsymbol{\Sigma}^{-1}\mathbf{S})\right\}\right]$
- ▶ Wishart Distribution: $W \sim W_d(n, \Omega)$ has the pdf

$$p(\mathbf{W}) \propto |\mathbf{W}|^{\frac{n-d-1}{2}} \exp\left\{-\frac{1}{2} \text{trace}(\mathbf{W} \mathbf{\Omega}^{-1})\right\}, \quad n>d-1, \; \mathbf{\Omega}>0, \; \mathbf{W} \in \mathbf{M}^+.$$

- ▶ If $z_1, \ldots, z_n \sim \text{Normal}(0, \sigma^2)$, then $\sum_{i=1}^n z_i^2 \sim \sigma^2 \chi_n^2$.
- ▶ If $\mathbf{z}_1, \dots, \mathbf{z}_n \sim \text{MVN}_d(\mathbf{0}, \mathbf{\Omega})$, then $\mathbf{Z}^T \mathbf{Z} \sim W_d(n, \mathbf{\Omega})$.
- ▶ If $\mathbf{W} \sim W_d(n, \mathbf{\Omega})$, then $\mathbb{E}(\mathbf{W}) = n\mathbf{\Omega}$.
- ▶ If $\mathbf{W} \sim W_d(n, \mathbf{\Omega})$, then, for any $\mathbf{B}^{q \times d}$, $\mathbf{B} \mathbf{W} \mathbf{B}^{\mathrm{T}} \sim W_q(n, \mathbf{B} \mathbf{\Omega} \mathbf{B}^{\mathrm{T}})$.
- ▶ If $\mathbf{W} = ((w_{ij})) \sim W_d(n, \mathbf{\Omega})$ with $\mathbf{\Omega} = ((\omega_{ij}))$, then $w_{ii} \sim \omega_{ii} \chi_n^2$.
- ▶ If $\mathbf{W}_j \stackrel{\text{that}}{\sim} W_d(n_j, \mathbf{\Omega})$, then $\sum_{j=1}^J \mathbf{W}_j \sim W_d(\sum_{j=1}^J n_j, \mathbf{\Omega})$.
- ▶ Inverse-Wishart Distribution: If $W \sim W_d(n, \Omega)$ and $V = W^{-1}$ and $\Lambda = \Omega^{-1}$, then $V \sim IW_d(n, \Lambda)$ and has the pdf

 $p(\mathbf{V}) \propto |\mathbf{V}|^{-\frac{N+\Omega+1}{2}} \exp\left\{-\frac{1}{2}\operatorname{trace}(\mathbf{V}^{-1}\mathbf{\Lambda})\right\}, \quad n > d-1, \ \mathbf{\Lambda} > 0, \ \mathbf{V} \in \mathbf{M}^+.$

11/18 33/71

$$\mathbf{y}_1, \dots, \mathbf{y}_n \overset{iid}{\sim} \text{MVN}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$
 with $\boldsymbol{\mu}^{d \times 1}, \boldsymbol{\Sigma}^{d \times d}$ both unknown

- ▶ MVN Likelihood: $p(\mathbf{y}_{1:n} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(\mathbf{y}_{i} \boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{y}_{i} \boldsymbol{\mu})\right\}$ $\propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left[-\frac{1}{2}\left\{n(\boldsymbol{\mu} - \overline{\mathbf{y}})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \overline{\mathbf{y}}) + \operatorname{trace}(\boldsymbol{\Sigma}^{-1}\mathbf{S})\right\}\right]$
- ▶ Wishart Distribution: $\mathbf{W} \sim W_d(n, \mathbf{\Omega})$ has the pdf $p(\mathbf{W}) \propto |\mathbf{W}|^{\frac{n-d-1}{2}} \exp\left\{-\frac{1}{2} \operatorname{trace}(\mathbf{W}\mathbf{\Omega}^{-1})\right\}, \quad n > d-1, \ \mathbf{\Omega} > 0, \ \mathbf{W} \in \mathbf{M}^+.$
 - ▶ If $z_1, \ldots, z_n \sim \text{Normal}(0, \sigma^2)$, then $\sum_{i=1}^n z_i^2 \sim \sigma^2 \chi_n^2$.
 - ▶ If $\mathbf{z}_1, \dots, \mathbf{z}_n \sim \text{MVN}_d(\mathbf{0}, \mathbf{\Omega})$, then $\mathbf{Z}^T \mathbf{Z} \sim W_d(n, \mathbf{\Omega})$.
 - ▶ If $\mathbf{W} \sim W_d(n, \mathbf{\Omega})$, then $\mathbb{E}(\mathbf{W}) = n\mathbf{\Omega}$.
 - ▶ If $\mathbf{W} \sim W_d(n, \mathbf{\Omega})$, then, for any $\mathbf{B}^{q \times d}$, $\mathbf{B}\mathbf{W}\mathbf{B}^{\mathrm{T}} \sim W_q(n, \mathbf{B}\mathbf{\Omega}\mathbf{B}^{\mathrm{T}})$.
 - ▶ If $\mathbf{W} = ((w_{ij})) \sim W_d(n, \Omega)$ with $\Omega = ((\omega_{ij}))$, then $w_{ii} \sim \omega_{ii} \chi_n^2$.
 - ▶ If $\mathbf{W}_j \stackrel{ind}{\sim} W_d(n_j, \Omega)$, then $\sum_{j=1}^J \mathbf{W}_j \sim W_d(\sum_{j=1}^J n_j, \Omega)$.
- ▶ Inverse-Wishart Distribution: If $W \sim W_d(n, \Omega)$ and $V = W^{-1}$ and $\Lambda = \Omega^{-1}$ then $V \sim IW_d(n, \Lambda)$ and has the pdf

 $p(\mathbf{V}) \propto |\mathbf{V}|^{-\frac{N+\Omega+1}{2}} \exp\left\{-\frac{1}{2}\operatorname{trace}(\mathbf{V}^{-1}\mathbf{\Lambda})\right\}, \quad n > d-1, \ \mathbf{\Lambda} > 0, \ \mathbf{V} \in \mathbf{M}^+.$

11/18 34/71

$$\mathbf{y}_1, \dots, \mathbf{y}_n \overset{iid}{\sim} \text{MVN}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$
 with $\boldsymbol{\mu}^{d \times 1}, \boldsymbol{\Sigma}^{d \times d}$ both unknown

- ▶ MVN Likelihood: $p(\mathbf{y}_{1:n} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(\mathbf{y}_{i} \boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{y}_{i} \boldsymbol{\mu})\right\}$ $\propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left[-\frac{1}{2}\left\{n(\boldsymbol{\mu} - \overline{\mathbf{y}})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \overline{\mathbf{y}}) + \operatorname{trace}(\boldsymbol{\Sigma}^{-1}\mathbf{S})\right\}\right]$
- ▶ Wishart Distribution: $\mathbf{W} \sim W_d(n, \mathbf{\Omega})$ has the pdf $p(\mathbf{W}) \propto |\mathbf{W}|^{\frac{n-d-1}{2}} \exp\left\{-\frac{1}{2} \operatorname{trace}(\mathbf{W}\mathbf{\Omega}^{-1})\right\}, \quad n > d-1, \ \mathbf{\Omega} > 0, \ \mathbf{W} \in \mathbf{M}^+.$
 - ▶ If $z_1, \ldots, z_n \sim \text{Normal}(0, \sigma^2)$, then $\sum_{i=1}^n z_i^2 \sim \sigma^2 \chi_n^2$.
 - ▶ If $\mathbf{z}_1, \dots, \mathbf{z}_n \sim \text{MVN}_d(\mathbf{0}, \mathbf{\Omega})$, then $\mathbf{Z}^T \mathbf{Z} \sim W_d(n, \mathbf{\Omega})$.
 - ▶ If $\mathbf{W} \sim W_d(n, \mathbf{\Omega})$, then $\mathbb{E}(\mathbf{W}) = n\mathbf{\Omega}$.
 - ▶ If $\mathbf{W} \sim W_d(n, \mathbf{\Omega})$, then, for any $\mathbf{B}^{q \times d}$, $\mathbf{B}\mathbf{W}\mathbf{B}^{\mathrm{T}} \sim W_q(n, \mathbf{B}\mathbf{\Omega}\mathbf{B}^{\mathrm{T}})$.
 - ▶ If $\mathbf{W} = ((w_{ij})) \sim W_d(n, \mathbf{\Omega})$ with $\mathbf{\Omega} = ((\omega_{ij}))$, then $w_{ii} \sim \omega_{ii} \chi_n^2$.
 - ▶ If $\mathbf{W}_j \stackrel{ind}{\sim} W_d(n_j, \mathbf{\Omega})$, then $\sum_{j=1}^J \mathbf{W}_j \sim W_d(\sum_{j=1}^J n_j, \mathbf{\Omega})$.
- ▶ Inverse-Wishart Distribution: If $W \sim W_d(n, \Omega)$ and $V = W^{-1}$ and $\Lambda = \Omega^{-1}$ then $V \sim IW_d(n, \Lambda)$ and has the pdf

 $p(\mathbf{V}) \propto |\mathbf{V}|^{-\frac{n+a+1}{2}} \exp\left\{-\frac{1}{2}\operatorname{trace}(\mathbf{V}^{-1}\mathbf{\Lambda})\right\}, \quad n > d-1, \ \mathbf{\Lambda} > 0, \ \mathbf{V} \in \mathbf{M}^+.$

11/18 35/71

$$\mathbf{y}_1, \dots, \mathbf{y}_n \overset{iid}{\sim} \text{MVN}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$
 with $\boldsymbol{\mu}^{d \times 1}, \boldsymbol{\Sigma}^{d \times d}$ both unknown

- ▶ MVN Likelihood: $p(\mathbf{y}_{1:n} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(\mathbf{y}_{i} \boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{y}_{i} \boldsymbol{\mu})\right\}$ $\propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left[-\frac{1}{2}\left\{n(\boldsymbol{\mu} - \overline{\mathbf{y}})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \overline{\mathbf{y}}) + \operatorname{trace}(\boldsymbol{\Sigma}^{-1}\mathbf{S})\right\}\right]$
- ▶ Wishart Distribution: $\mathbf{W} \sim W_d(n, \mathbf{\Omega})$ has the pdf $p(\mathbf{W}) \propto |\mathbf{W}|^{\frac{n-d-1}{2}} \exp\left\{-\frac{1}{2} \mathrm{trace}(\mathbf{W}\mathbf{\Omega}^{-1})\right\}, \quad n > d-1, \; \mathbf{\Omega} > 0, \; \mathbf{W} \in \mathbf{M}^+.$
 - ▶ If $z_1, \ldots, z_n \sim \text{Normal}(0, \sigma^2)$, then $\sum_{i=1}^n z_i^2 \sim \sigma^2 \chi_n^2$.
 - ▶ If $\mathbf{z}_1, \dots, \mathbf{z}_n \sim \text{MVN}_d(\mathbf{0}, \mathbf{\Omega})$, then $\mathbf{Z}^T \mathbf{Z} \sim W_d(n, \mathbf{\Omega})$.
 - ▶ If $\mathbf{W} \sim W_d(n, \mathbf{\Omega})$, then $\mathbb{E}(\mathbf{W}) = n\mathbf{\Omega}$.
 - ▶ If $\mathbf{W} \sim W_d(n, \mathbf{\Omega})$, then, for any $\mathbf{B}^{q \times d}$, $\mathbf{B} \mathbf{W} \mathbf{B}^{\mathrm{T}} \sim W_q(n, \mathbf{B} \mathbf{\Omega} \mathbf{B}^{\mathrm{T}})$.
 - ▶ If $\mathbf{W} = ((w_{ij})) \sim W_d(n, \mathbf{\Omega})$ with $\mathbf{\Omega} = ((\omega_{ij}))$, then $w_{ii} \sim \omega_{ii} \chi_n^2$.
 - ▶ If $\mathbf{W}_j \stackrel{ind}{\sim} W_d(n_j, \mathbf{\Omega})$, then $\sum_{j=1}^J \mathbf{W}_j \sim W_d(\sum_{j=1}^J n_j, \mathbf{\Omega})$.
- ▶ Inverse-Wishart Distribution: If $W \sim W_d(n, \Omega)$ and $V = W^{-1}$ and $\Lambda = \Omega^{-1}$ then $V \sim IW_d(n, \Lambda)$ and has the pdf

 $p(\mathbf{V}) \propto |\mathbf{V}|^{-\frac{N+\Omega+1}{2}} \exp\left\{-\frac{1}{2}\operatorname{trace}(\mathbf{V}^{-1}\mathbf{\Lambda})\right\}, \quad n > d-1, \ \mathbf{\Lambda} > 0, \ \mathbf{V} \in \mathbf{M}^+.$

11/18 36/71

Multivariate Normal Model Wishart and Inverse Wishart Distributions

$$\mathbf{y}_1, \dots, \mathbf{y}_n \overset{iid}{\sim} \text{MVN}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$
 with $\boldsymbol{\mu}^{d \times 1}, \boldsymbol{\Sigma}^{d \times d}$ both unknown

- ▶ MVN Likelihood: $p(\mathbf{y}_{1:n} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(\mathbf{y}_{i} \boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{y}_{i} \boldsymbol{\mu})\right\}$ $\propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left[-\frac{1}{2}\left\{n(\boldsymbol{\mu} - \overline{\mathbf{y}})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \overline{\mathbf{y}}) + \operatorname{trace}(\boldsymbol{\Sigma}^{-1}\mathbf{S})\right\}\right]$
- ▶ Wishart Distribution: $\mathbf{W} \sim W_d(n, \mathbf{\Omega})$ has the pdf $p(\mathbf{W}) \propto |\mathbf{W}|^{\frac{n-d-1}{2}} \exp\left\{-\frac{1}{2} \operatorname{trace}(\mathbf{W}\mathbf{\Omega}^{-1})\right\}, \quad n > d-1, \ \mathbf{\Omega} > 0, \ \mathbf{W} \in \mathbf{M}^+.$
 - ▶ If $z_1, \ldots, z_n \sim \text{Normal}(0, \sigma^2)$, then $\sum_{i=1}^n z_i^2 \sim \sigma^2 \chi_n^2$.
 - ▶ If $\mathbf{z}_1, \dots, \mathbf{z}_n \sim \text{MVN}_d(\mathbf{0}, \mathbf{\Omega})$, then $\mathbf{Z}^T \mathbf{Z} \sim W_d(n, \mathbf{\Omega})$.
 - ▶ If $\mathbf{W} \sim W_d(n, \mathbf{\Omega})$, then $\mathbb{E}(\mathbf{W}) = n\mathbf{\Omega}$.
 - ▶ If $\mathbf{W} \sim W_d(n, \mathbf{\Omega})$, then, for any $\mathbf{B}^{q \times d}$, $\mathbf{B}\mathbf{W}\mathbf{B}^{\mathrm{T}} \sim W_q(n, \mathbf{B}\mathbf{\Omega}\mathbf{B}^{\mathrm{T}})$.
 - ▶ If $\mathbf{W} = ((w_{ij})) \sim W_d(n, \Omega)$ with $\Omega = ((\omega_{ij}))$, then $w_{ii} \sim \omega_{ii} \chi_n^2$.
 - ▶ If $\mathbf{W}_j \stackrel{ind}{\sim} W_d(n_j, \mathbf{\Omega})$, then $\sum_{j=1}^J \mathbf{W}_j \sim W_d(\sum_{j=1}^J n_j, \mathbf{\Omega})$.
- ▶ Inverse-Wishart Distribution: If $\mathbf{W} \sim W_d(n, \mathbf{\Omega})$ and $\mathbf{V} = \mathbf{W}^{-1}$ and $\mathbf{\Lambda} = \mathbf{\Omega}^{-1}$, then $\mathbf{V} \sim \mathrm{IW}_d(n, \mathbf{\Lambda})$ and has the pdf

$$p(\mathbf{V}) \propto |\mathbf{V}|^{-\frac{n+d+1}{2}} \exp\left\{-\frac{1}{2}\mathrm{trace}(\mathbf{V}^{-1}\mathbf{\Lambda})\right\}, \quad n>d-1, \ \mathbf{\Lambda}>0, \ \mathbf{V} \in \mathbf{M}^+.$$

▶ If $\mathbf{V} \sim \mathrm{IW}_d(n, \mathbf{\Lambda})$, then $\mathbb{E}(\mathbf{V}) = \mathbf{\Lambda}/(n-d-1)$ provided n > d+1.

11/18 37/71

Multivariate Normal Model Wishart and Inverse Wishart Distributions

$$\mathbf{y}_1, \dots, \mathbf{y}_n \overset{iid}{\sim} \text{MVN}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$
 with $\boldsymbol{\mu}^{d \times 1}, \boldsymbol{\Sigma}^{d \times d}$ both unknown

- ▶ MVN Likelihood: $p(\mathbf{y}_{1:n} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(\mathbf{y}_{i} \boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{y}_{i} \boldsymbol{\mu})\right\}$ $\propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left[-\frac{1}{2}\left\{n(\boldsymbol{\mu} - \overline{\mathbf{y}})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \overline{\mathbf{y}}) + \operatorname{trace}(\boldsymbol{\Sigma}^{-1}\mathbf{S})\right\}\right]$
- ▶ Wishart Distribution: $\mathbf{W} \sim W_d(n, \mathbf{\Omega})$ has the pdf

$$p(\mathbf{W}) \propto |\mathbf{W}|^{\frac{n-d-1}{2}} \exp\left\{-\frac{1}{2} \mathrm{trace}(\mathbf{W} \mathbf{\Omega}^{-1})\right\}, \quad n > d-1, \; \mathbf{\Omega} > 0, \; \mathbf{W} \in \mathbf{M}^+.$$

- ▶ If $z_1, \ldots, z_n \sim \text{Normal}(0, \sigma^2)$, then $\sum_{i=1}^n z_i^2 \sim \sigma^2 \chi_n^2$.
- ▶ If $\mathbf{z}_1, \dots, \mathbf{z}_n \sim \text{MVN}_d(\mathbf{0}, \mathbf{\Omega})$, then $\mathbf{Z}^T \mathbf{Z} \sim W_d(n, \mathbf{\Omega})$.
- ▶ If $\mathbf{W} \sim W_d(n, \mathbf{\Omega})$, then $\mathbb{E}(\mathbf{W}) = n\mathbf{\Omega}$.
- ▶ If $\mathbf{W} \sim W_d(n, \mathbf{\Omega})$, then, for any $\mathbf{B}^{q \times d}$, $\mathbf{B}\mathbf{W}\mathbf{B}^{\mathrm{T}} \sim W_q(n, \mathbf{B}\mathbf{\Omega}\mathbf{B}^{\mathrm{T}})$.
- ▶ If $\mathbf{W} = ((w_{ij})) \sim W_d(n, \Omega)$ with $\Omega = ((\omega_{ij}))$, then $w_{ii} \sim \omega_{ii} \chi_n^2$.
- ▶ If $\mathbf{W}_j \stackrel{ind}{\sim} W_d(n_j, \mathbf{\Omega})$, then $\sum_{j=1}^J \mathbf{W}_j \sim W_d(\sum_{j=1}^J n_j, \mathbf{\Omega})$.
- ▶ Inverse-Wishart Distribution: If $\mathbf{W} \sim W_d(n, \mathbf{\Omega})$ and $\mathbf{V} = \mathbf{W}^{-1}$ and $\mathbf{\Lambda} = \mathbf{\Omega}^{-1}$, then $\mathbf{V} \sim \mathrm{IW}_d(n, \mathbf{\Lambda})$ and has the pdf
 - $p(\mathbf{V}) \propto |\mathbf{V}|^{-\frac{n+d+1}{2}} \exp\left\{-\frac{1}{2}\operatorname{trace}(\mathbf{V}^{-1}\mathbf{\Lambda})\right\}, \quad n > d-1, \ \mathbf{\Lambda} > 0, \ \mathbf{V} \in \mathbf{M}^+.$
 - ▶ If $\mathbf{V} \sim \mathrm{IW}_d(n, \mathbf{\Lambda})$, then $\mathbb{E}(\mathbf{V}) = \mathbf{\Lambda}/(n-d-1)$ provided n > d+1.

11/18 38/71

Multivariate Normal Model Conjugate Prior

$$\mathbf{y}_1, \dots, \mathbf{y}_n \overset{iid}{\sim} \text{MVN}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$
 with $\boldsymbol{\mu}^{d \times 1}, \boldsymbol{\Sigma}^{d \times d}$ both unknown

- ▶ MVN Likelihood: $p(\mathbf{y}_{1:n} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left\{-\frac{1}{2} \sum_{i=1}^{n} (\mathbf{y}_i \boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{y}_i \boldsymbol{\mu})\right\}$ $\propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left[-\frac{1}{2} \left\{n(\boldsymbol{\mu} - \overline{\mathbf{y}})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - \overline{\mathbf{y}}) + \operatorname{trace}(\boldsymbol{\Sigma}^{-1} \mathbf{S})\right\}\right]$
- ▶ MVN-Inverse-Wishart Prior: $(\mu, \Sigma) \sim \text{NIW}(\mu_0, \Sigma_0/\kappa_0, \nu_0, \Sigma_0)$

$$\begin{split} p(\pmb{\mu}, \pmb{\Sigma}) &= p(\pmb{\Sigma}) p(\pmb{\mu} \mid \pmb{\Sigma}) = \text{IW}\left(\pmb{\Sigma} \mid \nu_0, \pmb{\Sigma}_0\right) \cdot \text{MVN}(\pmb{\mu} \mid \pmb{\mu}_0, \pmb{\Sigma}/\kappa_0) \\ &\propto \left|\pmb{\Sigma}\right|^{-\left(\frac{\nu_0 + d + 1}{2} + \frac{1}{2}\right)} \exp\left[-\frac{1}{2}\left\{\text{trace}(\pmb{\Sigma}^{-1}\pmb{\Sigma}_0) + \kappa_0(\pmb{\mu} - \pmb{\mu}_0)^{\text{T}}\pmb{\Sigma}^{-1}(\pmb{\mu} - \pmb{\mu}_0)\right\}\right] \end{split}$$

MVN-Inverse-Wishart Posterior:

$$\begin{split} & p(\mu, \Sigma \mid \mathbf{y}_{1:n}) \propto |\Sigma|^{-\left\{\frac{\nu_0 + n + d + 1}{2} + \frac{1}{2}\right\}} \exp\left[-\frac{1}{2}\left\{\mathrm{trace}(\Sigma^{-1}\Sigma_n) + (\kappa_0 + n)(\mu - \mu_n)^{\mathrm{T}}\Sigma^{-1}(\mu - \mu_n)\right\}\right] \\ & \equiv \mathrm{NW}\left(\mu_n, \Sigma_n / \kappa_n, \nu_n, \Sigma_n\right), \quad \nu_n = (\nu_0 + n), \quad \kappa_n = (\kappa_0 + n), \quad \mu_n = (\kappa_0 \mu_0 + n \overline{\mathbf{y}}) / (\kappa_0 + n), \\ & \Sigma_n = \left\{\Sigma_0 + \mathbf{S} + \frac{n\kappa_0}{(n + \kappa_0)}(\overline{\mathbf{y}} - \mu_0)(\overline{\mathbf{y}} - \mu_0)^{\mathrm{T}}\right\}, \quad \mathbf{S} = \sum_{i=1}^n (\mathbf{y}_i - \overline{\mathbf{y}})(\mathbf{y}_i - \overline{\mathbf{y}})^{\mathrm{T}}. \end{split}$$

12/18 39/71

Multivariate Normal Model Conjugate Prior

$$\mathbf{y}_1, \dots, \mathbf{y}_n \overset{iid}{\sim} \text{MVN}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$
 with $\boldsymbol{\mu}^{d \times 1}, \boldsymbol{\Sigma}^{d \times d}$ both unknown

- ▶ MVN Likelihood: $p(\mathbf{y}_{1:n} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left\{-\frac{1}{2} \sum_{i=1}^{n} (\mathbf{y}_i \boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{y}_i \boldsymbol{\mu})\right\}$ $\propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left[-\frac{1}{2} \left\{n(\boldsymbol{\mu} - \overline{\mathbf{y}})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - \overline{\mathbf{y}}) + \operatorname{trace}(\boldsymbol{\Sigma}^{-1} \mathbf{S})\right\}\right]$
- ▶ MVN-Inverse-Wishart Prior: $(\mu, \Sigma) \sim \text{NIW}(\mu_0, \Sigma_0/\kappa_0, \nu_0, \Sigma_0)$

$$\begin{split} p(\pmb{\mu}, \pmb{\Sigma}) &= p(\pmb{\Sigma}) p(\pmb{\mu} \mid \pmb{\Sigma}) = \text{IW}\left(\pmb{\Sigma} \mid \nu_0, \pmb{\Sigma}_0\right) \cdot \text{MVN}(\pmb{\mu} \mid \pmb{\mu}_0, \pmb{\Sigma}/\kappa_0) \\ &\propto \left|\pmb{\Sigma}\right|^{-\left(\frac{\nu_0 + d + 1}{2} + \frac{1}{2}\right)} \exp\left[-\frac{1}{2}\left\{\text{trace}(\pmb{\Sigma}^{-1}\pmb{\Sigma}_0) + \kappa_0(\pmb{\mu} - \pmb{\mu}_0)^{\text{T}}\pmb{\Sigma}^{-1}(\pmb{\mu} - \pmb{\mu}_0)\right\}\right] \end{split}$$

► MVN-Inverse-Wishart Posterior:

$$\begin{split} & p(\boldsymbol{\mu}, \boldsymbol{\Sigma} \mid \mathbf{y}_{1:n}) \propto |\boldsymbol{\Sigma}| - \left\{ \frac{\nu_0 + n + d + 1}{2} + \frac{1}{2} \right\} \\ & \exp\left[-\frac{1}{2} \left\{ \operatorname{trace}\left[\boldsymbol{\Sigma}^{-1} \left\{ \boldsymbol{\Sigma}_0 + \mathbf{S} + \frac{n \kappa_0}{(n + \kappa_0)} (\overline{\mathbf{y}} - \boldsymbol{\mu}_0) (\overline{\mathbf{y}} - \boldsymbol{\mu}_0)^{\mathrm{T}} \right\} \right] + (\kappa_0 + n)(\boldsymbol{\mu} - \boldsymbol{\mu}_n) \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - \boldsymbol{\mu}_n)^{\mathrm{T}} \right\} \right] \\ & \propto |\boldsymbol{\Sigma}| - \left\{ \frac{\nu_0 + n + d + 1}{2} + \frac{1}{2} \right\} \exp\left[-\frac{1}{2} \left\{ \operatorname{trace}(\boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_n) + (\kappa_0 + n)(\boldsymbol{\mu} - \boldsymbol{\mu}_n)^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - \boldsymbol{\mu}_n) \right\} \right] \\ & \equiv \operatorname{NIW}(\boldsymbol{\mu}_n, \boldsymbol{\Sigma}_n / \kappa_n, \nu_n, \boldsymbol{\Sigma}_n) \,, \quad \nu_n = (\nu_0 + n), \quad \kappa_n = (\kappa_0 + n), \quad \boldsymbol{\mu}_n = (\kappa_0 \boldsymbol{\mu}_0 + n \overline{\mathbf{y}}) / (\kappa_0 + n), \\ & \boldsymbol{\Sigma}_n = \left\{ \boldsymbol{\Sigma}_0 + \mathbf{S} + \frac{n \kappa_0}{(n + \kappa_0)} (\overline{\mathbf{y}} - \boldsymbol{\mu}_0) (\overline{\mathbf{y}} - \boldsymbol{\mu}_0)^{\mathrm{T}} \right\}, \quad \mathbf{S} = \sum_{i=1}^n (\mathbf{y}_i - \overline{\mathbf{y}}) (\mathbf{y}_i - \overline{\mathbf{y}})^{\mathrm{T}}. \end{split}$$

12/18 40/71

Multivariate Normal Model Conjugate Prior

$$\mathbf{y}_1, \dots, \mathbf{y}_n \overset{iid}{\sim} \text{MVN}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$
 with $\boldsymbol{\mu}^{d \times 1}, \boldsymbol{\Sigma}^{d \times d}$ both unknown

- ▶ MVN Likelihood: $p(\mathbf{y}_{1:n} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left\{-\frac{1}{2} \sum_{i=1}^{n} (\mathbf{y}_i \boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{y}_i \boldsymbol{\mu})\right\}$ $\propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left[-\frac{1}{2} \left\{n(\boldsymbol{\mu} - \overline{\mathbf{y}})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - \overline{\mathbf{y}}) + \operatorname{trace}(\boldsymbol{\Sigma}^{-1} \mathbf{S})\right\}\right]$
- ▶ MVN-Inverse-Wishart Prior: $(\mu, \Sigma) \sim \text{NIW}(\mu_0, \Sigma_0/\kappa_0, \nu_0, \Sigma_0)$

$$\begin{split} p(\pmb{\mu}, \pmb{\Sigma}) &= p(\pmb{\Sigma}) p(\pmb{\mu} \mid \pmb{\Sigma}) = \text{IW}\left(\pmb{\Sigma} \mid \nu_0, \pmb{\Sigma}_0\right) \cdot \text{MVN}(\pmb{\mu} \mid \pmb{\mu}_0, \pmb{\Sigma}/\kappa_0) \\ &\propto \left|\pmb{\Sigma}\right|^{-\left(\frac{\nu_0 + d + 1}{2} + \frac{1}{2}\right)} \exp\left[-\frac{1}{2}\left\{\text{trace}(\pmb{\Sigma}^{-1}\pmb{\Sigma}_0) + \kappa_0(\pmb{\mu} - \pmb{\mu}_0)^{\text{T}}\pmb{\Sigma}^{-1}(\pmb{\mu} - \pmb{\mu}_0)\right\}\right] \end{split}$$

MVN-Inverse-Wishart Posterior:

$$\begin{split} & p(\boldsymbol{\mu}, \boldsymbol{\Sigma} \mid \mathbf{y}_{1:n}) \propto |\boldsymbol{\Sigma}| - \left\{ \frac{\nu_0 + n + d + 1}{2} + \frac{1}{2} \right\} \exp\left[-\frac{1}{2} \left\{ \operatorname{trace}(\boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_n) + (\kappa_0 + n)(\boldsymbol{\mu} - \boldsymbol{\mu}_n)^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - \boldsymbol{\mu}_n) \right\} \right] \\ & \equiv \operatorname{NIW}(\boldsymbol{\mu}_n, \boldsymbol{\Sigma}_n / \kappa_n, \boldsymbol{\nu}_n, \boldsymbol{\Sigma}_n) \,, \quad \boldsymbol{\nu}_n = (\nu_0 + n), \quad \kappa_n = (\kappa_0 + n), \quad \boldsymbol{\mu}_n = (\kappa_0 \boldsymbol{\mu}_0 + n \overline{\mathbf{y}}) / (\kappa_0 + n), \\ & \boldsymbol{\Sigma}_n = \left\{ \boldsymbol{\Sigma}_0 + \mathbf{S} + \frac{n \kappa_0}{(n + \kappa_0)} (\overline{\mathbf{y}} - \boldsymbol{\mu}_0) (\overline{\mathbf{y}} - \boldsymbol{\mu}_0)^T \right\}, \quad \mathbf{S} = \sum_{i=1}^n (\mathbf{y}_i - \overline{\mathbf{y}}) (\mathbf{y}_i - \overline{\mathbf{y}})^T. \end{split}$$

- **▶** Conditional Posteriors:

12/18 41/71

Multivariate Normal Model Improper Prior

$$\mathbf{y}_1, \dots, \mathbf{y}_n \overset{iid}{\sim} \text{MVN}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$
 with $\boldsymbol{\mu}^{d \times 1}, \boldsymbol{\Sigma}^{d \times d}$ both unknown

- ▶ MVN Likelihood: $p(\mathbf{y}_{1:n} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left\{-\frac{1}{2} \sum_{i=1}^{n} (\mathbf{y}_i \boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{y}_i \boldsymbol{\mu})\right\}$ $\propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left[-\frac{1}{2} \left\{n(\boldsymbol{\mu} - \overline{\mathbf{y}})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - \overline{\mathbf{y}}) + \operatorname{trace}(\boldsymbol{\Sigma}^{-1} \mathbf{S})\right\}\right]$
- ▶ Improper Prior: $p(\mu, \Sigma) \propto |\Sigma|^{-\frac{d+1}{2}}$
- ► Posterior:

$$\begin{split} &p(\boldsymbol{\mu}, \boldsymbol{\Sigma} \mid \mathbf{y}_{1:n}) \propto |\boldsymbol{\Sigma}|^{-\left\{\frac{n+d+1}{2}\right\}} \exp\left[-\frac{1}{2}\left\{\mathrm{trace}(\boldsymbol{\Sigma}^{-1}\mathbf{S}) + n(\boldsymbol{\mu} - \overline{\mathbf{y}})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \overline{\mathbf{y}})\right\}\right] \\ &= \text{NIW}\left(\overline{\mathbf{y}}, \mathbf{S}/n, n-1, \mathbf{S}\right) \end{split}$$

► Conditional Posteriors:

$$\triangleright p(\mu \mid \Sigma, \mathbf{y}_{1:n}) = \text{MVN}_d(\overline{\mathbf{y}}, \Sigma/n),$$

13/18 42/71

Multivariate Normal Model Improper Prior

$$\mathbf{y}_1, \dots, \mathbf{y}_n \overset{iid}{\sim} \text{MVN}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$
 with $\boldsymbol{\mu}^{d \times 1}, \boldsymbol{\Sigma}^{d \times d}$ both unknown

- ▶ MVN Likelihood: $p(\mathbf{y}_{1:n} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left\{-\frac{1}{2} \sum_{i=1}^{n} (\mathbf{y}_i \boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{y}_i \boldsymbol{\mu})\right\}$ $\propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left[-\frac{1}{2} \left\{n(\boldsymbol{\mu} - \overline{\mathbf{y}})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - \overline{\mathbf{y}}) + \operatorname{trace}(\boldsymbol{\Sigma}^{-1} \mathbf{S})\right\}\right]$
- ▶ Improper Prior: $p(\mu, \Sigma) \propto |\Sigma|^{-\frac{d+1}{2}}$
- **▶** Posterior:

$$p(\boldsymbol{\mu}, \boldsymbol{\Sigma} \mid \mathbf{y}_{1:n}) \propto |\boldsymbol{\Sigma}|^{-\left\{\frac{n+d+1}{2}\right\}} \exp\left[-\frac{1}{2}\left\{\operatorname{trace}(\boldsymbol{\Sigma}^{-1}\mathbf{S}) + n(\boldsymbol{\mu} - \overline{\mathbf{y}})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \overline{\mathbf{y}})\right\}\right]$$

$$\equiv \operatorname{NIW}\left(\overline{\mathbf{y}}, \mathbf{S}/n, n - 1, \mathbf{S}\right)$$

▶ Conditional Posteriors:

$$\triangleright p(\mu \mid \Sigma, \mathbf{y}_{1:n}) = MVN_d(\overline{\mathbf{y}}, \Sigma/n),$$

13/18 43/71

Multivariate Normal Model Improper Prior

$$\mathbf{y}_1, \dots, \mathbf{y}_n \overset{iid}{\sim} \text{MVN}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$
 with $\boldsymbol{\mu}^{d \times 1}, \boldsymbol{\Sigma}^{d \times d}$ both unknown

- ▶ MVN Likelihood: $p(\mathbf{y}_{1:n} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left\{-\frac{1}{2} \sum_{i=1}^{n} (\mathbf{y}_i \boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{y}_i \boldsymbol{\mu})\right\}$ $\propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left[-\frac{1}{2} \left\{n(\boldsymbol{\mu} - \overline{\mathbf{y}})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - \overline{\mathbf{y}}) + \operatorname{trace}(\boldsymbol{\Sigma}^{-1} \mathbf{S})\right\}\right]$
- ▶ Improper Prior: $p(\mu, \Sigma) \propto |\Sigma|^{-\frac{d+1}{2}}$
- **▶** Posterior:

$$p(\boldsymbol{\mu}, \boldsymbol{\Sigma} \mid \mathbf{y}_{1:n}) \propto |\boldsymbol{\Sigma}|^{-\left\{\frac{n+d+1}{2}\right\}} \exp\left[-\frac{1}{2}\left\{\operatorname{trace}(\boldsymbol{\Sigma}^{-1}\mathbf{S}) + n(\boldsymbol{\mu} - \overline{\mathbf{y}})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \overline{\mathbf{y}})\right\}\right]$$

$$\equiv \operatorname{NIW}\left(\overline{\mathbf{y}}, \mathbf{S}/n, n-1, \mathbf{S}\right)$$

- **▶** Conditional Posteriors:

13/18 44/71

Multivariate Normal Model Semi-Conjugate Prior

$$\mathbf{y}_1, \dots, \mathbf{y}_n \overset{iid}{\sim} \text{MVN}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$
 with $\boldsymbol{\mu}^{d \times 1}, \boldsymbol{\Sigma}^{d \times d}$ both unknown

- ► MVN Likelihood: $p(\mathbf{y}_{1:n} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left\{-\frac{1}{2} \sum_{i=1}^{n} (\mathbf{y}_i \boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{y}_i \boldsymbol{\mu})\right\}$ $\propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left[-\frac{1}{2} \left\{n(\boldsymbol{\mu} - \overline{\mathbf{y}})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - \overline{\mathbf{y}}) + \operatorname{trace}(\boldsymbol{\Sigma}^{-1} \mathbf{S})\right\}\right]$
- ▶ Prior: $p(\mu, \Sigma) = \text{MVN}(\mu \mid \mu_0, \Omega_0) \times \text{IW}(\Sigma \mid \nu_0, \Sigma_0)$
- ▶ Posterior:

$$\begin{split} & p(\boldsymbol{\mu}, \boldsymbol{\Sigma} \mid \mathbf{y}_{1:n}) \propto \exp\left\{-\frac{1}{2}(\boldsymbol{\mu} - \boldsymbol{\mu}_0)^{\mathrm{T}}\boldsymbol{\Omega}_0^{-1}(\boldsymbol{\mu} - \boldsymbol{\mu}_0)\right\} \left|\boldsymbol{\Sigma}\right|^{-\left\{\frac{\nu_0 + d}{2} + \frac{1}{2}\right\}} \exp\left\{-\frac{1}{2}\mathrm{trace}(\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_0)\right\} \\ & \times \left|\boldsymbol{\Sigma}\right|^{-\frac{n}{2}} \exp\left[-\frac{1}{2}\left\{\mathrm{trace}(\boldsymbol{\Sigma}^{-1}\mathbf{S}) + n(\boldsymbol{\mu} - \overline{\mathbf{y}})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \overline{\mathbf{y}})\right\}\right]. \end{split}$$

▶ Conditional Posteriors:

$$p(\mu \mid \Sigma, \mathbf{y}_{1:n}) = \text{MVN}_d(\mu_n, \Omega_n),$$

$$\mathbf{Q}^{-1} = (\mathbf{Q}^{-1} + n\Sigma^{-1}), \quad \mu = \mathbf{Q},$$

$$\mu_n = (a \iota_0 + i \iota_2), \quad \mu_n = a \iota_n (a \iota_0) \mu_0 + i \iota_2 \quad y),$$

14/18 45/71

Multivariate Normal Model Semi-Conjugate Prior

$$\mathbf{y}_1, \dots, \mathbf{y}_n \overset{iid}{\sim} \text{MVN}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$
 with $\boldsymbol{\mu}^{d \times 1}, \boldsymbol{\Sigma}^{d \times d}$ both unknown

- ▶ MVN Likelihood: $p(\mathbf{y}_{1:n} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(\mathbf{y}_{i} \boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{y}_{i} \boldsymbol{\mu})\right\}$ $\propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left[-\frac{1}{2}\left\{n(\boldsymbol{\mu} - \overline{\mathbf{y}})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \overline{\mathbf{y}}) + \operatorname{trace}(\boldsymbol{\Sigma}^{-1}\mathbf{S})\right\}\right]$
- ▶ Prior: $p(\mu, \Sigma) = MVN(\mu \mid \mu_0, \Omega_0) \times IW(\Sigma \mid \nu_0, \Sigma_0)$
- **▶** Posterior:

$$\begin{split} & p(\boldsymbol{\mu}, \boldsymbol{\Sigma} \mid \mathbf{y}_{1:n}) \propto \exp\left\{-\frac{1}{2}(\boldsymbol{\mu} - \boldsymbol{\mu}_0)^{\mathrm{T}} \boldsymbol{\Omega}_0^{-1}(\boldsymbol{\mu} - \boldsymbol{\mu}_0)\right\} |\boldsymbol{\Sigma}|^{-\left\{\frac{\nu_0 + d}{2} + \frac{1}{2}\right\}} \exp\left\{-\frac{1}{2} \mathrm{trace}(\boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_0)\right\} \\ & \times |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left[-\frac{1}{2} \left\{ \mathrm{trace}(\boldsymbol{\Sigma}^{-1} \mathbf{S}) + n(\boldsymbol{\mu} - \overline{\mathbf{y}})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \overline{\mathbf{y}})\right\}\right]. \end{split}$$

► Conditional Posteriors:

$$\blacktriangleright p(\mu \mid \Sigma, \mathbf{y}_{1:n}) = \text{MVN}_d(\mu_n, \Omega_n),$$

$$\Omega_n^{-1} = (\Omega_0^{-1} + n\Sigma^{-1}), \quad \mu_n = \Omega_n(\Omega_0^{-1}\mu_0 + n\Sigma^{-1}\overline{y}).$$

14/18 46/71

Multivariate Normal Model Semi-Conjugate Prior

$$\mathbf{y}_1, \dots, \mathbf{y}_n \overset{iid}{\sim} \text{MVN}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$
 with $\boldsymbol{\mu}^{d \times 1}, \boldsymbol{\Sigma}^{d \times d}$ both unknown

- ▶ MVN Likelihood: $p(\mathbf{y}_{1:n} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left\{-\frac{1}{2} \sum_{i=1}^{n} (\mathbf{y}_i \boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{y}_i \boldsymbol{\mu})\right\}$ $\propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left[-\frac{1}{2} \left\{n(\boldsymbol{\mu} - \overline{\mathbf{y}})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - \overline{\mathbf{y}}) + \operatorname{trace}(\boldsymbol{\Sigma}^{-1} \mathbf{S})\right\}\right]$
- ▶ Prior: $p(\mu, \Sigma) = MVN(\mu \mid \mu_0, \Omega_0) \times IW(\Sigma \mid \nu_0, \Sigma_0)$
- **Posterior:**

$$\begin{split} & p(\boldsymbol{\mu}, \boldsymbol{\Sigma} \mid \mathbf{y}_{1:n}) \propto \exp\left\{-\frac{1}{2}(\boldsymbol{\mu} - \boldsymbol{\mu}_0)^{\mathrm{T}} \boldsymbol{\Omega}_0^{-1}(\boldsymbol{\mu} - \boldsymbol{\mu}_0)\right\} |\boldsymbol{\Sigma}|^{-\left\{\frac{\nu_0 + d}{2} + \frac{1}{2}\right\}} \exp\left\{-\frac{1}{2} \mathrm{trace}(\boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_0)\right\} \\ & \times |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left[-\frac{1}{2} \left\{ \mathrm{trace}(\boldsymbol{\Sigma}^{-1} \mathbf{S}) + n(\boldsymbol{\mu} - \overline{\mathbf{y}})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \overline{\mathbf{y}})\right\}\right]. \end{split}$$

▶ Conditional Posteriors:

$$\boldsymbol{\Omega}_n^{-1} = (\boldsymbol{\Omega}_0^{-1} + n\boldsymbol{\Sigma}^{-1}), \ \boldsymbol{\mu}_n = \boldsymbol{\Omega}_n (\boldsymbol{\Omega}_0^{-1} \boldsymbol{\mu}_0 + n\boldsymbol{\Sigma}^{-1} \overline{\mathbf{y}}),$$

14/18 47/71

$$\mathbf{y}_1, \dots, \mathbf{y}_n \overset{iid}{\sim} \operatorname{Mult}(m, \pi_1, \dots, \pi_K)$$

with $\sum_{k=1}^K \pi_k = 1$ and $\mathbf{y}_i^{\mathrm{T}} = (y_{i1}, \dots, y_{iK})^{K \times 1}$ for all i

▶ Multinomial Likelihood: $p(\mathbf{y}_{1:n} \mid \boldsymbol{\pi}) = \prod_{i=1}^{n} \left\{ \frac{m!}{y_{i1}! \dots y_{iK}!} \pi_1^{y_{i1}} \dots \pi_K^{y_{iK}} \right\}$ $\propto \pi_1^{\sum_{i=1}^{n} y_{i1}} \dots \pi_K^{\sum_{i=1}^{n} y_{iK}} = \pi_1^{S_1} \dots \pi_K^{S_K}$

▶ Dirichlet Distributions: $\boldsymbol{\pi} = (\pi_1, \dots, \pi_K)^T \sim \text{Dir}(\alpha_1, \dots, \alpha_K)$ has the pdf

$$p(\boldsymbol{\pi}) = \frac{\Gamma(\alpha_1 + \dots + \alpha_K)}{\Gamma(\alpha_1) \dots \Gamma(\alpha_K)} \pi_1^{\alpha_1 - 1} \dots \pi_K^{\alpha_K - 1}, \quad \sum_{i=1}^K \pi_k = 1, \quad \alpha_k > 0 \ \forall \ k$$

15/18 48/71

$$\mathbf{y}_1, \dots, \mathbf{y}_n \overset{iid}{\sim} \operatorname{Mult}(m, \pi_1, \dots, \pi_K)$$

with $\sum_{k=1}^K \pi_k = 1$ and $\mathbf{y}_i^{\mathrm{T}} = (y_{i1}, \dots, y_{iK})^{K \times 1}$ for all i

▶ Multinomial Likelihood: $p(\mathbf{y}_{1:n} \mid \boldsymbol{\pi}) = \prod_{i=1}^n \left\{ \frac{m!}{y_{i1}! \dots y_{iK}!} \pi_1^{y_{i1}} \dots \pi_K^{y_{iK}} \right\}$ $\propto \pi_1^{\sum_{i=1}^n y_{i1}} \dots \pi_K^{\sum_{i=1}^n y_{iK}} = \pi_1^{S_1} \dots \pi_K^{S_K}$

▶ **Dirichlet Distributions:** $\boldsymbol{\pi} = (\pi_1, \dots, \pi_K)^T \sim \text{Dir}(\alpha_1, \dots, \alpha_K)$ has the pdf $p(\boldsymbol{\pi}) = \frac{\Gamma(\alpha_1 + \dots + \alpha_K)}{\Gamma(\alpha_K)} \frac{\pi_1^{\alpha_1 - 1}}{\Gamma(\alpha_K)} \dots \frac{\pi_K^{\alpha_K - 1}}{\Gamma(\alpha_K)}, \quad \sum_{i=1}^K \pi_k = 1, \quad \alpha_k > 0 \ \forall \ k.$

$$ightharpoonup \operatorname{Dir}(\alpha_1, \alpha_2) \equiv \operatorname{Beta}(\alpha_1, \alpha_2).$$

$$\mathbb{E}(\pi_k) = \frac{\alpha_k}{\alpha}, \quad \operatorname{var}(\pi_k) = \frac{\alpha_k(\alpha - \alpha_k)}{\alpha^2(\alpha + 1)}, \quad \alpha = \alpha_1 + \dots + \alpha_K$$

▶ If $z_k \stackrel{ind}{\sim} \text{Ga}(\alpha_k, 1)$ and $\pi_k \propto z_k$, then $\pi \sim \text{Dir}(\alpha_1, \dots, \alpha_K)$. \rightarrow important result

▶ If
$$\pi \sim \text{Dir}(\alpha_1, \ldots, \alpha_K)$$
, then

$$(\pi_1 + \dots + \pi_j, \pi_{j+1}, \dots, \pi_K)^{\perp} \sim \operatorname{Dir}(\alpha_1 + \dots + \alpha_j, \alpha_{j+1}, \dots, \alpha_K)$$

▶ If $\pi \sim \text{Dir}(\alpha_1, \dots, \alpha_K)$, then $\left(\frac{\pi_1}{\sum_{j=1}^J \pi_j}, \dots, \frac{\pi_J}{\sum_{j=1}^J \pi_j}\right)^T \sim \text{Dir}(\alpha_1, \dots, \alpha_J)$.

15/18 49/71

$$\mathbf{y}_1, \dots, \mathbf{y}_n \overset{iid}{\sim} \operatorname{Mult}(m, \pi_1, \dots, \pi_K)$$

with $\sum_{k=1}^K \pi_k = 1$ and $\mathbf{y}_i^{\mathrm{T}} = (y_{i1}, \dots, y_{iK})^{K \times 1}$ for all i

▶ Multinomial Likelihood: $p(\mathbf{y}_{1:n} \mid \boldsymbol{\pi}) = \prod_{i=1}^{n} \left\{ \frac{m!}{y_{i1}!...y_{iK}!} \pi_{1}^{y_{i1}} \dots \pi_{K}^{y_{iK}} \right\}$ $\propto \pi_{1}^{\sum_{i=1}^{n} y_{i1}} \dots \pi_{K}^{\sum_{i=1}^{n} y_{iK}} = \pi_{1}^{S_{1}} \dots \pi_{K}^{S_{K}}$

- ▶ **Dirichlet Distributions:** $\boldsymbol{\pi} = (\pi_1, \dots, \pi_K)^{\mathrm{T}} \sim \mathrm{Dir}(\alpha_1, \dots, \alpha_K)$ has the pdf $p(\boldsymbol{\pi}) = \frac{\Gamma(\alpha_1 + \dots + \alpha_K)}{\Gamma(\alpha_1) \dots \Gamma(\alpha_K)} \pi_1^{\alpha_1 1} \dots \pi_K^{\alpha_K 1}, \quad \sum_{i=1}^K \pi_k = 1, \quad \alpha_k > 0 \ \forall \ k.$

 - $\mathbb{E}(\pi_k) = \frac{\alpha_k}{\alpha}, \quad \text{var}(\pi_k) = \frac{\alpha_k(\alpha \alpha_k)}{\alpha^2(\alpha + 1)}, \quad \alpha = \alpha_1 + \dots + \alpha_K.$
 - ▶ If $z_k \stackrel{ind}{\sim} \operatorname{Ga}(\alpha_k, 1)$ and $\pi_k \propto z_k$, then $\pi \sim \operatorname{Dir}(\alpha_1, \dots, \alpha_K)$. \to important resulting
 - If $\pi \sim \text{Dir}(\alpha_1, \dots, \alpha_K)$, then $(\pi_1 + \dots + \pi_j, \pi_{j+1}, \dots, \pi_K)^{\mathrm{T}} \sim \text{Dir}(\alpha_1 + \dots + \alpha_j, \alpha_{j+1}, \dots, \alpha_K)$
 - ▶ If $\pi \sim \text{Dir}(\alpha_1, \dots, \alpha_K)$, then $\left(\frac{\pi_1}{\sum_{j=1}^J \pi_j}, \dots, \frac{\pi_J}{\sum_{j=1}^J \pi_j}\right)^1 \sim \text{Dir}(\alpha_1, \dots, \alpha_J)$.

15/18 50/71

$$\mathbf{y}_1, \dots, \mathbf{y}_n \overset{iid}{\sim} \operatorname{Mult}(m, \pi_1, \dots, \pi_K)$$

with $\sum_{k=1}^K \pi_k = 1$ and $\mathbf{y}_i^{\mathrm{T}} = (y_{i1}, \dots, y_{iK})^{K \times 1}$ for all i

▶ Multinomial Likelihood: $p(\mathbf{y}_{1:n} \mid \pi) = \prod_{i=1}^n \left\{ \frac{m!}{y_{i1}!...y_{iK}!} \pi_1^{y_{i1}} \dots \pi_K^{y_{iK}} \right\}$ $\propto \pi_1^{\sum_{i=1}^n y_{i1}} \dots \pi_K^{\sum_{i=1}^n y_{iK}} = \pi_1^{S_1} \dots \pi_K^{S_K}$

- ▶ **Dirichlet Distributions:** $\boldsymbol{\pi} = (\pi_1, \dots, \pi_K)^{\mathrm{T}} \sim \mathrm{Dir}(\alpha_1, \dots, \alpha_K)$ has the pdf $p(\boldsymbol{\pi}) = \frac{\Gamma(\alpha_1 + \dots + \alpha_K)}{\Gamma(\alpha_1) \dots \Gamma(\alpha_K)} \pi_1^{\alpha_1 1} \dots \pi_K^{\alpha_K 1}, \quad \sum_{i=1}^K \pi_k = 1, \quad \alpha_k > 0 \ \forall \ k.$

 - $\blacktriangleright \mathbb{E}(\pi_k) = \frac{\alpha_k}{\alpha}, \quad \operatorname{var}(\pi_k) = \frac{\alpha_k(\alpha \alpha_k)}{\alpha^2(\alpha + 1)}, \quad \alpha = \alpha_1 + \dots + \alpha_K.$
 - ▶ If $z_k \stackrel{ina}{\sim} \operatorname{Ga}(\alpha_k, 1)$ and $\pi_k \propto z_k$, then $\pi \sim \operatorname{Dir}(\alpha_1, \dots, \alpha_K)$. \to important result!
 - ▶ If $\pi \sim \text{Dir}(\alpha_1, \dots, \alpha_K)$, then $(\pi_1 + \dots + \pi_2, \pi_{2+1}, \dots, \pi_K)^T \sim \text{Dir}(\alpha_1 + \dots + \alpha_2, \alpha_{2+1})$
 - $(\pi_1 + \dots + \pi_j, \pi_{j+1}, \dots, \pi_K)^{\perp} \sim \operatorname{Dir}(\alpha_1 + \dots + \alpha_j, \alpha_{j+1}, \dots, \alpha_K)$
 - ▶ If $\pi \sim \text{Dir}(\alpha_1, \dots, \alpha_K)$, then $\left(\frac{\pi_1}{\sum_{j=1}^J \pi_j}, \dots, \frac{\pi_J}{\sum_{j=1}^J \pi_j}\right)^T \sim \text{Dir}(\alpha_1, \dots, \alpha_J)$.

15/18 51/71

$$\mathbf{y}_1, \dots, \mathbf{y}_n \overset{iid}{\sim} \operatorname{Mult}(m, \pi_1, \dots, \pi_K)$$

with $\sum_{k=1}^K \pi_k = 1$ and $\mathbf{y}_i^{\mathrm{T}} = (y_{i1}, \dots, y_{iK})^{K \times 1}$ for all i

- ▶ Multinomial Likelihood: $p(\mathbf{y}_{1:n} \mid \pi) = \prod_{i=1}^n \left\{ \frac{m!}{y_{i1}!...y_{iK}!} \pi_1^{y_{i1}} \dots \pi_K^{y_{iK}} \right\}$ $\propto \pi_1^{\sum_{i=1}^n y_{i1}} \dots \pi_K^{\sum_{i=1}^n y_{iK}} = \pi_1^{S_1} \dots \pi_K^{S_K}$
- ▶ **Dirichlet Distributions:** $\boldsymbol{\pi} = (\pi_1, \dots, \pi_K)^T \sim \text{Dir}(\alpha_1, \dots, \alpha_K)$ has the pdf $p(\boldsymbol{\pi}) = \frac{\Gamma(\alpha_1 + \dots + \alpha_K)}{\Gamma(\alpha_1) \dots \Gamma(\alpha_K)} \pi_1^{\alpha_1 1} \dots \pi_K^{\alpha_K 1}, \quad \sum_{i=1}^K \pi_k = 1, \quad \alpha_k > 0 \ \forall \ k.$

 - $\blacktriangleright \mathbb{E}(\pi_k) = \frac{\alpha_k}{\alpha}, \quad \text{var}(\pi_k) = \frac{\alpha_k(\alpha \alpha_k)}{\alpha^2(\alpha + 1)}, \quad \alpha = \alpha_1 + \dots + \alpha_K.$
 - ▶ If $z_k \stackrel{ind}{\sim} \text{Ga}(\alpha_k, 1)$ and $\pi_k \propto z_k$, then $\pi \sim \text{Dir}(\alpha_1, \dots, \alpha_K)$. \rightarrow important result!
 - If $\pi \sim \text{Dir}(\alpha_1, \dots, \alpha_K)$, then $(\pi_1 + \dots + \pi_j, \pi_{j+1}, \dots, \pi_K)^T \sim \text{Dir}(\alpha_1 + \dots + \alpha_j, \alpha_{j+1}, \dots, \alpha_K)$
 - ▶ If $\pi \sim \text{Dir}(\alpha_1, \dots, \alpha_K)$, then $\left(\frac{\pi_1}{\sum_{j=1}^J \pi_j}, \dots, \frac{\pi_J}{\sum_{j=1}^J \pi_j}\right)^T \sim \text{Dir}(\alpha_1, \dots, \alpha_J)$.

15/18 52/71

$$\mathbf{y}_1, \dots, \mathbf{y}_n \overset{iid}{\sim} \operatorname{Mult}(m, \pi_1, \dots, \pi_K)$$

with $\sum_{k=1}^K \pi_k = 1$ and $\mathbf{y}_i^{\mathrm{T}} = (y_{i1}, \dots, y_{iK})^{K \times 1}$ for all i

- ▶ Multinomial Likelihood: $p(\mathbf{y}_{1:n} \mid \pi) = \prod_{i=1}^n \left\{ \frac{m!}{y_{i1}!...y_{iK}!} \pi_1^{y_{i1}} \dots \pi_K^{y_{iK}} \right\}$ $\propto \pi_1^{\sum_{i=1}^n y_{i1}} \dots \pi_K^{\sum_{i=1}^n y_{iK}} = \pi_1^{S_1} \dots \pi_K^{S_K}$
- ▶ **Dirichlet Distributions:** $\boldsymbol{\pi} = (\pi_1, \dots, \pi_K)^{\mathrm{T}} \sim \mathrm{Dir}(\alpha_1, \dots, \alpha_K)$ has the pdf $p(\boldsymbol{\pi}) = \frac{\Gamma(\alpha_1 + \dots + \alpha_K)}{\Gamma(\alpha_1) \dots \Gamma(\alpha_K)} \pi_1^{\alpha_1 1} \dots \pi_K^{\alpha_K 1}, \quad \sum_{i=1}^K \pi_k = 1, \quad \alpha_k > 0 \ \forall \ k.$

 - $\mathbb{E}(\pi_k) = \frac{\alpha_k}{\alpha}, \quad \operatorname{var}(\pi_k) = \frac{\alpha_k(\alpha \alpha_k)}{\alpha^2(\alpha + 1)}, \quad \alpha = \alpha_1 + \dots + \alpha_K.$
 - ▶ If $z_k \stackrel{ind}{\sim} \text{Ga}(\alpha_k, 1)$ and $\pi_k \propto z_k$, then $\pi \sim \text{Dir}(\alpha_1, \dots, \alpha_K)$. \rightarrow important result!
 - ► If $\pi \sim \text{Dir}(\alpha_1, \dots, \alpha_K)$, then $(\pi_1 + \dots + \pi_j, \pi_{j+1}, \dots, \pi_K)^T \sim \text{Dir}(\alpha_1 + \dots + \alpha_j, \alpha_{j+1}, \dots, \alpha_K)$.

 $If \pi \sim Dir(\alpha_1, \dots, \alpha_K), then \left(\frac{\pi_1}{\sum_{j=1}^J \pi_j}, \dots, \frac{\pi_J}{\sum_{j=1}^J \pi_j}\right)^1 \sim Dir(\alpha_1, \dots, \alpha_J)$

15/18 53/71

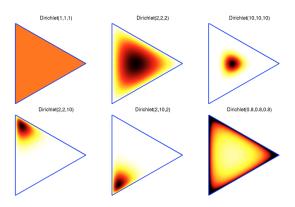
$$\mathbf{y}_1, \dots, \mathbf{y}_n \overset{iid}{\sim} \operatorname{Mult}(m, \pi_1, \dots, \pi_K)$$

with $\sum_{k=1}^K \pi_k = 1$ and $\mathbf{y}_i^{\mathrm{T}} = (y_{i1}, \dots, y_{iK})^{K \times 1}$ for all i

- ▶ Multinomial Likelihood: $p(\mathbf{y}_{1:n} \mid \pi) = \prod_{i=1}^n \left\{ \frac{m!}{y_{i1}!...y_{iK}!} \pi_1^{y_{i1}} \dots \pi_K^{y_{iK}} \right\}$ $\propto \pi_1^{\sum_{i=1}^n y_{i1}} \dots \pi_K^{\sum_{i=1}^n y_{iK}} = \pi_1^{S_1} \dots \pi_K^{S_K}$
- ▶ **Dirichlet Distributions:** $\boldsymbol{\pi} = (\pi_1, \dots, \pi_K)^T \sim \text{Dir}(\alpha_1, \dots, \alpha_K)$ has the pdf $p(\boldsymbol{\pi}) = \frac{\Gamma(\alpha_1 + \dots + \alpha_K)}{\Gamma(\alpha_1) \dots \Gamma(\alpha_K)} \pi_1^{\alpha_1 1} \dots \pi_K^{\alpha_K 1}, \quad \sum_{i=1}^K \pi_k = 1, \quad \alpha_k > 0 \ \forall \ k.$

 - $\mathbb{E}(\pi_k) = \frac{\alpha_k}{\alpha}, \quad \text{var}(\pi_k) = \frac{\alpha_k(\alpha \alpha_k)}{\alpha^2(\alpha + 1)}, \quad \alpha = \alpha_1 + \dots + \alpha_K.$
 - ▶ If $z_k \stackrel{ind}{\sim} \text{Ga}(\alpha_k, 1)$ and $\pi_k \propto z_k$, then $\pi \sim \text{Dir}(\alpha_1, \dots, \alpha_K)$. \rightarrow important result!
 - ▶ If $\pi \sim \text{Dir}(\alpha_1, \dots, \alpha_K)$, then $(\pi_1 + \dots + \pi_j, \pi_{j+1}, \dots, \pi_K)^T \sim \text{Dir}(\alpha_1 + \dots + \alpha_j, \alpha_{j+1}, \dots, \alpha_K)$.
 - ▶ If $\pi \sim \text{Dir}(\alpha_1, \dots, \alpha_K)$, then $\left(\frac{\pi_1}{\sum_{j=1}^J \pi_j}, \dots, \frac{\pi_J}{\sum_{j=1}^J \pi_j}\right)^T \sim \text{Dir}(\alpha_1, \dots, \alpha_J)$.

15/18 54/71



16/18 55/71

$$\begin{aligned} \mathbf{y}_1, \dots, \mathbf{y}_n &\overset{iid}{\sim} \mathsf{Mult}(m, \pi_1, \dots, \pi_K) \\ & \text{with } \pi_k \geq 0 \ \forall k, \ \sum_{k=1}^K \pi_k = 1 \ \text{and} \ \mathbf{y}_i^\mathrm{T} = (y_{i1}, \dots, y_{iK})^{K \times 1} \ \text{for all} \ i \end{aligned}$$

- ▶ Multinomial Likelihood: $p(\mathbf{y}_{1:n} \mid \boldsymbol{\pi}) = \prod_{i=1}^n \left\{ \frac{m!}{y_{i1}! \dots y_{iK}!} \pi_1^{y_{i1}} \dots \pi_K^{y_{iK}} \right\}$ $\propto \pi_1^{\sum_{i=1}^n y_{i1}} \dots \pi_K^{\sum_{i=1}^n y_{iK}} = \pi_1^{S_1} \dots \pi_K^{S_K}$
- **▶ Dirichlet Prior:** $\pi \sim \text{Dir}(\alpha_1, \dots, \alpha_K)$ has the pdf

$$p(\boldsymbol{\pi}) = \frac{\Gamma(\alpha_1 + \dots + \alpha_K)}{\Gamma(\alpha_1) \dots \Gamma(\alpha_K)} \pi_1^{\alpha_1 - 1} \dots \pi_K^{\alpha_K - 1}, \quad \pi_k \ge 0 \forall k, \ \sum_{i=1}^K \pi_k = 1, \ \alpha_k > 0 \ \forall \ k.$$

- Posterior: $p(\pi \mid \mathbf{y}_{1:n}) \propto \pi_1^{\alpha_1 + S_1 1} \dots \pi_K^{\alpha_K + S_K 1} \equiv \text{Dir}(\alpha_1 + S_1, \dots, \alpha_K + S_K)$
- ▶ **Jeffreys' NIP:** $\pi \sim \text{Dir}(1/2,...,1/2)$ so that $p(\pi) \propto \prod_{k=1}^K \pi_k^{-1/2}$.

17/18 56/71

$$\begin{aligned} \mathbf{y}_1, \dots, \mathbf{y}_n &\overset{iid}{\sim} \mathsf{Mult}(m, \pi_1, \dots, \pi_K) \\ & \text{with } \pi_k \geq 0 \ \forall k, \ \sum_{k=1}^K \pi_k = 1 \ \text{and} \ \mathbf{y}_i^\mathrm{T} = (y_{i1}, \dots, y_{iK})^{K \times 1} \ \text{for all} \ i \end{aligned}$$

- ▶ Multinomial Likelihood: $p(\mathbf{y}_{1:n} \mid \boldsymbol{\pi}) = \prod_{i=1}^n \left\{ \frac{m!}{y_{i1}! \dots y_{iK}!} \pi_1^{y_{i1}} \dots \pi_K^{y_{iK}} \right\}$ $\propto \pi_1^{\sum_{i=1}^n y_{i1}} \dots \pi_K^{\sum_{i=1}^n y_{iK}} = \pi_1^{S_1} \dots \pi_K^{S_K}$
- ▶ **Dirichlet Prior:** $\pi \sim \text{Dir}(\alpha_1, \dots, \alpha_K)$ has the pdf

$$p(\boldsymbol{\pi}) = \frac{\Gamma(\alpha_1 + \dots + \alpha_K)}{\Gamma(\alpha_1) \dots \Gamma(\alpha_K)} \pi_1^{\alpha_1 - 1} \dots \pi_K^{\alpha_K - 1}, \quad \pi_k \ge 0 \forall k, \ \sum_{i=1}^K \pi_k = 1, \ \alpha_k > 0 \ \forall \ k.$$

- ▶ Posterior: $p(\boldsymbol{\pi} \mid \mathbf{y}_{1:n}) \propto \pi_1^{\alpha_1 + S_1 1} \dots \pi_K^{\alpha_K + S_K 1} \equiv \text{Dir}(\alpha_1 + S_1, \dots, \alpha_K + S_K)$
- ▶ **Jeffreys' NIP:** $\pi \sim \text{Dir}(1/2, ..., 1/2)$ so that $p(\pi) \propto \prod_{k=1}^K \pi_k^{-1/2}$.

17/18 57/71

$$\begin{aligned} \mathbf{y}_1, \dots, \mathbf{y}_n &\overset{iid}{\sim} \mathsf{Mult}(m, \pi_1, \dots, \pi_K) \\ & \text{with } \pi_k \geq 0 \ \forall k, \ \sum_{k=1}^K \pi_k = 1 \ \text{and} \ \mathbf{y}_i^\mathrm{T} = (y_{i1}, \dots, y_{iK})^{K \times 1} \ \text{for all} \ i \end{aligned}$$

- ▶ Multinomial Likelihood: $p(\mathbf{y}_{1:n} \mid \boldsymbol{\pi}) = \prod_{i=1}^n \left\{ \frac{m!}{y_{i1}! \dots y_{iK}!} \pi_1^{y_{i1}} \dots \pi_K^{y_{iK}} \right\}$ $\propto \pi_1^{\sum_{i=1}^n y_{i1}} \dots \pi_K^{\sum_{i=1}^n y_{iK}} = \pi_1^{S_1} \dots \pi_K^{S_K}$
- **Dirichlet Prior:** $\pi \sim \text{Dir}(\alpha_1, \dots, \alpha_K)$ has the pdf

$$p(\boldsymbol{\pi}) = \frac{\Gamma(\alpha_1 + \dots + \alpha_K)}{\Gamma(\alpha_1) \dots \Gamma(\alpha_K)} \pi_1^{\alpha_1 - 1} \dots \pi_K^{\alpha_K - 1}, \quad \pi_k \ge 0 \forall k, \ \sum_{i=1}^K \pi_k = 1, \ \alpha_k > 0 \ \forall \ k.$$

- ▶ Posterior: $p(\boldsymbol{\pi} \mid \mathbf{y}_{1:n}) \propto \pi_1^{\alpha_1 + S_1 1} \dots \pi_K^{\alpha_K + S_K 1} \equiv \text{Dir}\left(\alpha_1 + S_1, \dots, \alpha_K + S_K\right)$
- ▶ **Jeffreys' NIP:** $\pi \sim \text{Dir}(1/2, ..., 1/2)$ so that $p(\pi) \propto \prod_{k=1}^K \pi_k^{-1/2}$.

17/18 58/71

$$\mathbf{y}_1, \dots, \mathbf{y}_n \overset{iid}{\sim} \mathsf{Mult}(m, \pi_1, \dots, \pi_K)$$
 with $\pi_k \geq 0 \ \forall k, \ \sum_{k=1}^K \pi_k = 1 \ \mathsf{and} \ \mathbf{y}_i^\mathrm{T} = (y_{i1}, \dots, y_{iK})^{K \times 1}$ for all i

- ▶ Multinomial Likelihood: $p(\mathbf{y}_{1:n} \mid \boldsymbol{\pi}) = \prod_{i=1}^n \left\{ \frac{m!}{y_{i1}! \dots y_{iK}!} \pi_1^{y_{i1}} \dots \pi_K^{y_{iK}} \right\}$ $\propto \pi_1^{\sum_{i=1}^n y_{i1}} \dots \pi_K^{\sum_{i=1}^n y_{iK}} = \pi_1^{S_1} \dots \pi_K^{S_K}$
- **Dirichlet Prior:** $\pi \sim \text{Dir}(\alpha_1, \dots, \alpha_K)$ has the pdf

$$p(\boldsymbol{\pi}) = \frac{\Gamma(\alpha_1 + \dots + \alpha_K)}{\Gamma(\alpha_1) \dots \Gamma(\alpha_K)} \pi_1^{\alpha_1 - 1} \dots \pi_K^{\alpha_K - 1}, \quad \pi_k \ge 0 \forall k, \ \sum_{i=1}^K \pi_k = 1, \ \alpha_k > 0 \ \forall \ k.$$

- ▶ Posterior: $p(\boldsymbol{\pi} \mid \mathbf{y}_{1:n}) \propto \pi_1^{\alpha_1 + S_1 1} \dots \pi_K^{\alpha_K + S_K 1} \equiv \text{Dir} (\alpha_1 + S_1, \dots, \alpha_K + S_K)$
- ▶ Jeffreys' NIP: $\pi \sim \text{Dir}(1/2, \dots, 1/2)$ so that $p(\pi) \propto \prod_{k=1}^K \pi_k^{-1/2}$.

17/18 59/71

$$\mathbf{y}_1, \dots, \mathbf{y}_n \overset{iid}{\sim} \mathsf{Mult}(m, \pi_1, \dots, \pi_K)$$
 with $\pi_k \geq 0 \ \forall k, \ \sum_{k=1}^K \pi_k = 1 \ \mathsf{and} \ \mathbf{y}_i^\mathrm{T} = (y_{i1}, \dots, y_{iK})^{K \times 1}$ for all i

- ▶ Multinomial Likelihood: $p(\mathbf{y}_{1:n} \mid \boldsymbol{\pi}) = \prod_{i=1}^n \left\{ \frac{m!}{y_{i1}! \dots y_{iK}!} \pi_1^{y_{i1}} \dots \pi_K^{y_{iK}} \right\}$ $\propto \pi_1^{\sum_{i=1}^n y_{i1}} \dots \pi_K^{\sum_{i=1}^n y_{iK}} = \pi_1^{S_1} \dots \pi_K^{S_K}$
- **▶ Dirichlet Prior:** $\pi \sim \text{Dir}(\alpha_1, \dots, \alpha_K)$ has the pdf

$$p(\boldsymbol{\pi}) = \frac{\Gamma(\alpha_1 + \dots + \alpha_K)}{\Gamma(\alpha_1) \dots \Gamma(\alpha_K)} \pi_1^{\alpha_1 - 1} \dots \pi_K^{\alpha_K - 1}, \quad \pi_k \ge 0 \forall k, \ \sum_{i=1}^K \pi_k = 1, \ \alpha_k > 0 \ \forall \ k.$$

- ▶ Posterior: $p(\boldsymbol{\pi} \mid \mathbf{y}_{1:n}) \propto \pi_1^{\alpha_1 + S_1 1} \dots \pi_K^{\alpha_K + S_K 1} \equiv \text{Dir}\left(\alpha_1 + S_1, \dots, \alpha_K + S_K\right)$
- ▶ **Jeffreys' NIP:** $\pi \sim \text{Dir}(1/2, ..., 1/2)$ so that $p(\pi) \propto \prod_{k=1}^K \pi_k^{-1/2}$.
- ▶ Uniform NIP: $\pi \sim \text{Dir}(1, ..., 1)$ so that $p(\pi) \propto 1$.

17/18 60/71

$$\begin{aligned} \mathbf{y}_1, \dots, \mathbf{y}_n &\overset{iid}{\sim} \mathsf{Mult}(m, \pi_1, \dots, \pi_K) \\ & \text{with } \pi_k \geq 0 \ \forall k, \ \sum_{k=1}^K \pi_k = 1 \ \text{and} \ \mathbf{y}_i^\mathrm{T} = (y_{i1}, \dots, y_{iK})^{K \times 1} \ \text{for all} \ i \end{aligned}$$

- ▶ Multinomial Likelihood: $p(\mathbf{y}_{1:n} \mid \boldsymbol{\pi}) = \prod_{i=1}^n \left\{ \frac{m!}{y_{i1}! \dots y_{iK}!} \pi_1^{y_{i1}} \dots \pi_K^{y_{iK}} \right\}$ $\propto \pi_1^{\sum_{i=1}^n y_{i1}} \dots \pi_K^{\sum_{i=1}^n y_{iK}} = \pi_1^{S_1} \dots \pi_K^{S_K}$
- **▶ Dirichlet Prior:** $\pi \sim \text{Dir}(\alpha_1, \dots, \alpha_K)$ has the pdf

$$p(\boldsymbol{\pi}) = \frac{\Gamma(\alpha_1 + \dots + \alpha_K)}{\Gamma(\alpha_1) \dots \Gamma(\alpha_K)} \pi_1^{\alpha_1 - 1} \dots \pi_K^{\alpha_K - 1}, \quad \pi_k \ge 0 \forall k, \ \sum_{i=1}^K \pi_k = 1, \ \alpha_k > 0 \ \forall k.$$

- ▶ Posterior: $p(\boldsymbol{\pi} \mid \mathbf{y}_{1:n}) \propto \pi_1^{\alpha_1 + S_1 1} \dots \pi_K^{\alpha_K + S_K 1} \equiv \text{Dir}\left(\alpha_1 + S_1, \dots, \alpha_K + S_K\right)$
- ▶ **Jeffreys' NIP:** $\pi \sim \text{Dir}(1/2, ..., 1/2)$ so that $p(\pi) \propto \prod_{k=1}^K \pi_k^{-1/2}$.
- ▶ Uniform NIP: $\pi \sim \text{Dir}(1, ..., 1)$ so that $p(\pi) \propto 1$.
 - ▶ In this case, $\widehat{\pi}_{MAP} = \widehat{\pi}_{MLE}$.

17/18 61/71

Normal model

- Normal-inverse-Gamma prior
- Improper prior
- Independent Normal-inverse-Gamma prior
- Multivariate normal mode
 - Normal inverse Wishart prior
 Improper prior
 Indicate Normal inverse Wishart
- Multinomial model

18/18 62/71

- Normal model
 - Normal-inverse-Gamma prior
 - Improper prio
 - Independent Normal-inverse-Gamma prior
- Multivariate normal model
- * Normal-inverse-Wishart prior
 - Ingraper prior
- Multinomial model

18/18 63/71

- Normal model
 - Normal-inverse-Gamma prior
 - Improper prior
 - Independent Normal-inverse-Gamma prior
- Multivariate normal model
- Normal-inverse-Wishart prior
 - Indexed the Normal-Toverse Wisham or a migraphen in t
- Multinomial model

18/18 64/71

- Normal model
 - Normal-inverse-Gamma prior
 - Improper prior
 - Independent Normal-inverse-Gamma prior
- Multivariate normal mode.
- * Independent Normal-inverse-Wishart prior
- Multinomial model

18/18 65/71

- Normal model
 - Normal-inverse-Gamma prior
 - Improper prior
 - Independent Normal-inverse-Gamma prior
- Multivariate normal model
 - Normal-inverse-Wishart prior
 - Improper prior
 - Independent Normal-inverse-Wishart prior
- Multinomial model

18/18 66/71

- Normal model
 - Normal-inverse-Gamma prior
 - Improper prior
 - Independent Normal-inverse-Gamma prior
- Multivariate normal model
 - Normal-inverse-Wishart prior
 - Improper prior
 - Independent Normal-inverse-Wishart prior
- Multinomial model

18/18 67/71

- Normal model
 - · Normal-inverse-Gamma prior
 - Improper prior
 - Independent Normal-inverse-Gamma prior
- Multivariate normal model
 - Normal-inverse-Wishart prior
 - Improper prior
 - Independent Normal-inverse-Wishart prior
- Multinomial model

18/18 68/71

- Normal model
 - · Normal-inverse-Gamma prior
 - Improper prior
 - Independent Normal-inverse-Gamma prior
- Multivariate normal model
 - Normal-inverse-Wishart prior
 - Improper prior
 - Independent Normal-inverse-Wishart prior

Multinomial model

18/18 69/71

- Normal model
 - · Normal-inverse-Gamma prior
 - Improper prior
 - Independent Normal-inverse-Gamma prior
- Multivariate normal model
 - Normal-inverse-Wishart prior
 - Improper prior
 - Independent Normal-inverse-Wishart prior
- Multinomial model

18/18 70/71

- Normal model
 - Normal-inverse-Gamma prior
 - Improper prior
 - Independent Normal-inverse-Gamma prior
- Multivariate normal model
 - Normal-inverse-Wishart prior
 - Improper prior
 - Independent Normal-inverse-Wishart prior
- Multinomial model
 - Dirichlet prior

18/18 71/71