1. Scale mixture of normal dist: If
$$\theta|\mu,\sigma^2,k \sim \mathrm{N}(\mu, \frac{\sigma^2}{k}), \quad k \sim \mathrm{G}(\frac{\nu}{2}, \frac{\nu}{2}),$$
then
$$\theta \sim t_\nu(\mu, \sigma^2)$$
2. Beta model
$$y|\theta \sim \mathrm{Binom}(n,\theta)$$
$$\theta \sim \mathrm{Beta}(\alpha,\beta) \quad (\text{Jeffrey: } \alpha=\beta=\frac{1}{2})$$
$$\theta|y \sim \mathrm{Beta}(\alpha+y, \beta+n-y)$$
3. Normal model (Known variance, unknown mean)
$$\pi(y|\mu) = (2\pi\sigma^2)^{-n/2}\exp\{-(2\sigma^2)^{-1}\sum_{i=1}^n (y_i-\mu)^2\}$$
$$\mu \sim \mathrm{N}(\zeta, \tau^2), \quad (\text{Jeffrey: } \pi(\mu) \propto 1)$$
$$\mu|y \sim \mathrm{N}(\zeta', \tau'^2)$$
$$\zeta' = (n\sigma^{-2}+\tau^{-2})^{-1}(n\sigma^{-2}\bar{y}+\tau^{-2}\zeta)$$
$$\tau'^2 = (n\sigma^{-2}+\tau^{-2})^{-1}$$
$$\tilde{y}|y \sim \mathrm{N}(\zeta', \sigma^2+\tau'^2)$$
4. Normal model (Known mean, unknown variance)
$$y|\sigma^2 \propto (\sigma^2)^{-\frac{n}{2}}\exp\{-\frac{1}{2}\sum (y_i-\mu)^2(\sigma^2)^{-1}\}$$
$$\sigma^2 \sim \mathrm{G}^{-1}(\alpha,\beta) \quad (\text{Jeffrey: } \alpha=1,\ \beta=0)$$
$$\sigma^2|y \sim \mathrm{G}^{-1}(\frac{n}{2}+\alpha, \frac{1}{2}\sum (y_i-\mu)^2+\beta)$$
5. Poisson model
$$\pi(y|\theta) \propto \theta^y \exp\{-\theta\}$$
$$\theta \sim \mathrm{G}(\alpha,\beta), \quad \theta|y \sim \mathrm{G}(\alpha+n\bar{y}, \beta+n)$$
$$\tilde{y} \sim \text{Neg-Binom}(\alpha,\beta)$$
$\tilde{y}$: # failures, α: # success until stop, β: odds of success
6. Jeffrey's prior
$$\pi(\theta) \propto [I(\theta)]^{\frac{1}{2}}, \quad I(\theta) = -E[\frac{d^2}{d\theta^2}\log \pi(y|\theta)]$$
7. Noninformative priors for pivital quantities
   (a) Location: $\pi(y-\theta|\theta) = f(y-\theta) \Rightarrow \pi(\theta) \propto 1$.
   (b) Scale: $\pi(y/\theta|\theta) = f(y/\theta) \Rightarrow \pi(\theta) \propto \theta^{-1}$
8. Normal model
(Unknown mean and variance, noninformative prior)
$$y_1,\ldots,y_n|\mu,\sigma^2 \overset{\text{iid}}{\sim} \mathrm{N}(\mu,\sigma^2)\pi(\mu,\sigma^{-2}) \propto (\sigma^{-2})^{-1}$$
$$\pi(\mu,\sigma^{-2}|y) \propto (\sigma^{-2})^{\frac{n}{2}-1}\exp\{-\frac{\sigma^{-2}}{2}[(n-1)s^2+n(\bar{y}-\mu)^2]\}$$
$$\sigma^{-2}|y \sim \mathrm{G}(\frac{n-1}{2}, \frac{(n-1)s^2}{2}), \quad \mu|\sigma^{-2},y \sim \mathrm{N}(\bar{y}, \frac{\sigma^2}{n})$$
$$\mu|y \sim t_{n-1}(\bar{y}, \frac{s}{\sqrt{n}})$$
$$\tilde{y}|y \sim t_{n-1}(\bar{y}, (1+\frac{1}{n})s^2) \ (\text{Scale mix norm, } k=\frac{s^2}{\sigma^2})$$
9. Normal model
(Unknown mean and variance, conjugate prior)
$$\sigma^{-2} \sim \mathrm{G}(\frac{\nu}{2}, \frac{\nu\tau^2}{2}), \quad \mu|\sigma^{-2} \sim \mathrm{N}(\theta, \frac{\sigma^2}{k})$$
$$\sigma^{-2}|y \sim \mathrm{G}(\frac{\nu+n}{2}, \frac{\nu\tau^2+(n-1)s^2+\frac{kn}{k+n}(\bar{y}-\theta)^2}{2})$$
$$\mu|\sigma^{-2},y \sim \mathrm{N}(\frac{k\theta+n\bar{y}}{k+n}, \frac{\sigma^2}{k+n})$$
$$\mu|y \sim t_{\nu+k}\left(\frac{k\theta+n\bar{y}}{k+n}, \frac{\nu\tau^2+(n-1)s^2+\frac{kn}{k+n}(\bar{y}-\theta)^2}{(\nu+n)(k+n)}\right)$$
10. Multinomial model
$$y|\theta \sim \mathrm{Multinom}(\theta), \quad \theta=(\theta_1,\ldots,\theta_n)^T, \quad \sum\theta_i = 1$$
$$\theta \sim \mathrm{Dirichlet}(\alpha_1,\ldots,\alpha_n), \quad \theta|y \sim \mathrm{Dirichlet}(\alpha_1+y_1,\ldots,\alpha_n+y_n)$$
11. Multivariate Normal
(Σ known, μ unknown, Conjugate prior)
$$y_i|\mu,\Sigma \overset{\text{iid}}{\sim} \mathrm{MVN}(\mu,\Sigma), \quad \underset{d\times n}{Y} = (y_1,\ldots,y_n)$$
$$\mu \sim \mathrm{MVN}(\theta,\Lambda), \quad \mu|Y \sim \mathrm{MVN}(\mu_p, \Lambda_p)$$
$$\mu_p = (\Lambda^{-1}+n\Sigma^{-1})^{-1}(\Lambda^{-1}\theta+n\Sigma^{-1}\bar{y})$$
$$\Lambda_p = (\Lambda^{-1}+n\Sigma^{-1})^{-1}$$
$$\mu_1|\mu_2,Y \sim \mathrm{MVN}(\mu_{p,1}+\Lambda_{p,12}\Lambda_{p,22}^{-1}(\mu_2-\mu_{p,2}), \Lambda_{p,1|2})$$
$$\Lambda_{p,1|2} = \Lambda_{p,11}-\Lambda_{p,12}\Lambda_{p,22}^{-1}\Lambda_{p,21}$$
$$\mu_1|Y \sim \mathrm{MVN}(\mu_{p,1}, \Lambda_{p,11})$$
$$\tilde{y}|Y \sim \mathrm{MVN}(\mu_p, \Lambda_p+\Sigma)$$
12. Multivariate Normal
(Σ known, μ unknown, Noninformative prior)
$$\pi(\mu) \propto 1, \quad \mu|Y \sim \mathrm{MVN}(\bar{y}, \frac{\Sigma}{n})$$
13. Multivariate Normal
(Σ, μ unknown, Conjugate prior)
$$\Sigma^{-1} \sim \mathrm{W}(\nu,\Lambda), \quad \mu|\Sigma^{-1} \sim \mathrm{N}(\theta, \frac{\Sigma}{k})$$
$$\Sigma^{-1}|Y \sim \mathrm{W}(\nu_p, \Lambda_p)$$
$$\nu_p = n+\nu, \quad \Lambda_p = \left(\Lambda^{-1}+S+\frac{kn}{k+n}(\theta-\bar{y})(\theta-\bar{y})^T\right)^{-1}$$
$$\mu|Y,\Sigma^{-1} \sim \mathrm{MVN}\left(\frac{k}{k+n}\theta+\frac{n}{k+n}\bar{y}, \frac{\Sigma}{k+n}\right)$$
$$\mu|Y \sim t_{\nu+n-d+1}\left(\frac{k\theta+n\bar{y}}{k+n}, \frac{\Lambda_p^{-1}}{(k+n)(\nu+n-d+1)}\right)$$
$$\tilde{y}|Y \sim t_{v+n-d+1}\left(\frac{k\theta+n\bar{y}}{k+n}, \left(1+\frac{1}{k+n}\right)\frac{\Lambda_p^{-1}}{v+n-d+1}\right)$$
14. Multivariate Normal
(Σ, μ unknown, Noninformative prior)
$$\pi(\mu,\Sigma) \propto |\Sigma|^{\frac{-(d+1)}{2}} \quad (\text{let } k\to 0, \nu\to -1, \Lambda^{-1}\to 0)$$
$$\Sigma^{-1}|Y \sim \mathrm{W}(n-1, S^{-1}), \quad \mu|Y,\Sigma^{-1} \sim \mathrm{MVN}(\bar{y}, \frac{\Sigma}{n})$$
$$\mu|Y \sim t_{n-d}\left(\bar{y}, \frac{S}{n(n-d)}\right)$$
$$\tilde{y}|Y \sim t_{n-d}\left(\bar{y}, \left(1+\frac{1}{n}\right)\frac{S}{n-d}\right), \quad S=\sum_{i=1}^n (y_i-\bar{y})(y_i-\bar{y})^T$$
15. Wishart distribution: If $z_1,\ldots,z_\nu \sim \mathrm{MVN}(0,\Lambda)$, then $\Sigma^{-1} = \sum_{l=1}^\nu z_l z_l^T \sim \mathrm{W}(\nu,\Lambda)$ and $\Sigma \sim W^{-1}(\nu,\Lambda^{-1})$.
16. Normal approximation
$$\pi(\theta|y) \approx \mathrm{N}(\hat{\boldsymbol{\theta}}^\pi, I^\pi(y)^{-1})$$
$$I_{ij}^\pi(y) = -[\frac{\partial^2}{\partial\theta_i\partial\theta_j}\log\{\pi(y|\boldsymbol{\theta})\pi(\boldsymbol{\theta})\}]_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}^\pi}$$
$$\hat{\boldsymbol{\theta}}^\pi = \text{posterior mode}$$
17. Laplace's method
$$I = \int f(\boldsymbol{\theta})\exp\{-nh(\boldsymbol{\theta})\}d\boldsymbol{\theta} \quad (\exp\{-nh(\boldsymbol{\theta})\} \propto \text{posterior})$$
f is smooth, positive; h is smooth, has unique min at $\boldsymbol{\theta}$.
$$\hat{I} = f(\hat{\boldsymbol{\theta}})(\frac{2\pi}{n})^{\frac{m}{2}}|\hat{\boldsymbol{\Sigma}}|^{\frac{1}{2}}\exp\{-nh(\hat{\boldsymbol{\theta}})\}\{1+\mathcal{O}(n^{-1})\}$$
$$\hat{\boldsymbol{\Sigma}} = [D^2h(\hat{\boldsymbol{\theta}})]^{-1}$$
$$E[g(\boldsymbol{\theta})] = g(\hat{\boldsymbol{\theta}})+\{1+\mathcal{O}(n^{-1})\}$$
$$E[g(\boldsymbol{\theta})] = \frac{\int\exp\{\log[g(\boldsymbol{\theta})]-nh(\boldsymbol{\theta})\}d\boldsymbol{\theta}}{\int\exp\{-nh(\boldsymbol{\theta})\}d\boldsymbol{\theta}} = \frac{\int\exp\{-nh^*(\boldsymbol{\theta})\}d\boldsymbol{\theta}}{\int\exp\{-nh(\boldsymbol{\theta})\}d\boldsymbol{\theta}}$$
$$= \frac{|\tilde{\boldsymbol{\Sigma}}^*|^{\frac{1}{2}}\exp\{-nh^*(\tilde{\boldsymbol{\theta}})\}}{|\hat{\boldsymbol{\Sigma}}|^{\frac{1}{2}}\exp\{-nh^*(\hat{\boldsymbol{\theta}})\}}\{1+\mathcal{O}(n^{-2})\}$$
$$\hat{\pi}(\theta_1|y) \propto |\hat{\boldsymbol{\Sigma}}(\theta_1)|^{\frac{1}{2}}\exp\{-nh(\theta_1, \hat{\boldsymbol{\theta}}_{\mathbf{2}}(\theta_1))\}$$
$$h(\theta_1, \boldsymbol{\theta_2}) = -\frac{1}{n}\log\{\pi(y|\theta_1,\boldsymbol{\theta_2})\pi(\theta_1,\boldsymbol{\theta_2})\}$$
18. Direct sampling: Draw samples to simulate population
$$p := P[a<\theta<b|\boldsymbol{y}]$$
$$\hat{p} = \frac{1}{N}\sum_{j=1}^N I_{(a,b)}(\theta_j)$$
$$\hat{p} = \frac{1}{Nh_N}\sum_{j=1}^N K\left(\frac{\theta-\theta_j}{h_N}\right), \ h_N\to 0, Nh_N\to\infty \text{ as } N\to\infty$$
19. Importance sampling
$$E[h(\theta)|y] = \frac{\int h(\theta)w(\theta)g(\theta)d\theta}{\int w(\theta)g(\theta)d\theta} \approx \frac{\sum_j h(\theta_j)w(\theta_j)}{\sum_j w(\theta_j)}$$
where $g(\theta) \approx \pi(\theta|y)$, $\theta_j \sim g(\theta)$, $w(\theta) = \pi(y|\theta)\pi(\theta)/g(\theta)$