Likelihood	Model parameters	Conjugate prior distribution	Prior hyperparameters	Posterior hyperparameters[note 1]	Interpretation of hyperparameters	Posterior predictive[note 2]
Bernoulli	p (probability)	Beta	$\alpha,\, \beta$	$\alpha + \sum_{i=1}^n x_i,\, \beta + n - \sum_{i=1}^n x_i$	$\alpha - 1$ successes, $\beta - 1$ failures[note 3]	$p(\tilde{x}=1) = \frac{\alpha'}{\alpha'+\beta'}$
Binomial	p (probability)	Beta	$\alpha,\, \beta$	$\alpha + \sum_{i=1}^n x_i,\, \beta + \sum_{i=1}^n N_i - \sum_{i=1}^n x_i$	$\alpha - 1$ successes, $\beta - 1$ failures[note 3]	$\operatorname{BetaBin}(\tilde{x}\mid\alpha',\beta')$ (beta-binomial)
Negative binomial with known failure number, r	p (probability)	Beta	$\alpha,\, \beta$	$\alpha + \sum_{i=1}^n x_i,\, \beta + rn$	$\alpha - 1$ total successes, $\beta - 1$ failures[note 3] (i.e., $\frac{\beta-1}{r}$ experiments, assuming r stays fixed)	$\operatorname{BetaNegBin}(\tilde{x}\mid\alpha',\beta')$ (beta-negative binomial)
Poisson	λ (rate)	Gamma	$k,\, \theta$	$k + \sum_{i=1}^n x_i,\, \frac{\theta}{n\theta+1}$	k total occurrences in $\frac{1}{\theta}$ intervals	$\operatorname{NB}(\tilde{x}\mid k',\theta')$ (negative binomial)
			$\alpha,\, \beta$[note 4]	$\alpha + \sum_{i=1}^n x_i,\, \beta + n$	α total occurrences in β intervals	$\operatorname{NB}\left(\tilde{x}\mid\alpha',\frac{1}{1+\beta'}\right)$ (negative binomial)
Categorical	$\boldsymbol{p}$ (probability vector), k (number of categories; i.e., size of $\boldsymbol{p}$)	Dirichlet	$\boldsymbol{\alpha}$	$\boldsymbol{\alpha} + (c_1,\ldots,c_k)$, where c_i is the number of observations in category i	$\alpha_i - 1$ occurrences of category i[note 3]	$p(\tilde{x}=i) = \frac{\alpha_i'}{\sum_i \alpha_i'} = \frac{\alpha_i + c_i}{\sum_i \alpha_i + n}$
Multinomial	$\boldsymbol{p}$ (probability vector), k (number of categories; i.e., size of $\boldsymbol{p}$)	Dirichlet	$\boldsymbol{\alpha}$	$\boldsymbol{\alpha} + \sum_{i=1}^n \mathbf{x}_i$	$\alpha_i - 1$ occurrences of category i[note 3]	$\operatorname{DirMult}(\tilde{\mathbf{x}}\mid\boldsymbol{\alpha}')$ (Dirichlet-multinomial)
Hypergeometric with known total population size, N	M (number of target members)	Beta-binomial[4]	$n = N, \alpha,\, \beta$	$\alpha + \sum_{i=1}^n x_i,\, \beta + \sum_{i=1}^n N_i - \sum_{i=1}^n x_i$	$\alpha - 1$ successes, $\beta - 1$ failures[note 3]	
Geometric	p_0 (probability)	Beta	$\alpha,\, \beta$	$\alpha + n,\, \beta + \sum_{i=1}^n x_i$	$\alpha - 1$ experiments, $\beta - 1$ total failures[note 3]	

Likelihood	Model parameters	Conjugate prior distribution	Prior hyperparameters	Posterior hyperparameters[note 1]	Interpretation of hyperparameters	Posterior predictive[note 5]
Normal with known variance σ^2	μ (mean)	Normal	$\mu_0,\, \sigma_0^2$	$\frac{1}{\frac{1}{\sigma_0^2}+\frac{n}{\sigma^2}}\left(\frac{\mu_0}{\sigma_0^2}+\frac{\sum_{i=1}^n x_i}{\sigma^2}\right),\left(\frac{1}{\sigma_0^2}+\frac{n}{\sigma^2}\right)^{-1}$	mean was estimated from observations with total precision (sum of all individual precisions) $1/\sigma_0^2$ and with sample mean μ_0	$\mathcal{N}(\tilde{x}\vert\mu_0', {\sigma_0^2}' + \sigma^2)$[5]
Normal with known precision τ	μ (mean)	Normal	$\mu_0,\, \tau_0$	$\frac{\tau_0\mu_0 + \tau\sum_{i=1}^n x_i}{\tau_0 + n\tau},\, \tau_0 + n\tau$	mean was estimated from observations with total precision (sum of all individual precisions) τ_0 and with sample mean μ_0	$\mathcal{N}\left(\tilde{x}\mid\mu_0', \frac{1}{\tau_0'}+\frac{1}{\tau}\right)$[5]
Normal with known mean μ	σ^2 (variance)	Inverse gamma	$\alpha,\, \beta$ [note 6]	$\alpha+\frac{n}{2},\, \beta+\frac{\sum_{i=1}^n (x_i-\mu)^2}{2}$	variance was estimated from 2α observations with sample variance β/α (i.e. with sum of squared deviations 2β, where deviations are from known mean μ)	$t_{2\alpha'}(\tilde{x}\vert\mu,\sigma^2 = \beta'/\alpha')$[5]
Normal with known mean μ	σ^2 (variance)	Scaled inverse chi-squared	$\nu,\, \sigma_0^2$	$\nu+n,\, \frac{\nu\sigma_0^2 + \sum_{i=1}^n (x_i-\mu)^2}{\nu+n}$	variance was estimated from ν observations with sample variance σ_0^2	$t_{\nu'}(\tilde{x}\vert\mu,{\sigma_0^2}')$[5]
Normal with known mean μ	τ (precision)	Gamma	$\alpha,\, \beta$[note 4]	$\alpha+\frac{n}{2},\, \beta+\frac{\sum_{i=1}^n (x_i-\mu)^2}{2}$	precision was estimated from 2α observations with sample variance β/α (i.e. with sum of squared deviations 2β, where deviations are from known mean μ)	$t_{2\alpha'}(\tilde{x}\mid\mu,\sigma^2 = \beta'/\alpha')$[5]
Normal[note 7]	μ and σ^2 Assuming exchangeability	Normal-inverse gamma	$\mu_0,\, \nu,\, \alpha,\, \beta$	$\frac{\nu\mu_0+n\bar{x}}{\nu+n},\, \nu+n,\, \alpha+\frac{n}{2},$ $\beta+\tfrac{1}{2}\sum_{i=1}^n (x_i-\bar{x})^2 + \frac{n\nu}{\nu+n}\frac{(\bar{x}-\mu_0)^2}{2}$ • $\bar{x}$ is the sample mean	mean was estimated from ν observations with sample mean μ_0; variance was estimated from 2α observations with sample mean μ_0 and sum of squared deviations 2β	$t_{2\alpha'}\left(\tilde{x}\mid\mu', \frac{\beta'(\nu'+1)}{\nu'\alpha'}\right)$[5]
Normal	μ and τ Assuming exchangeability	Normal-gamma	$\mu_0,\, \nu,\, \alpha,\, \beta$	$\frac{\nu\mu_0+n\bar{x}}{\nu+n},\, \nu+n,\, \alpha+\frac{n}{2},$ $\beta+\tfrac{1}{2}\sum_{i=1}^n (x_i-\bar{x})^2 + \frac{n\nu}{\nu+n}\frac{(\bar{x}-\mu_0)^2}{2}$ • $\bar{x}$ is the sample mean	mean was estimated from ν observations with sample mean μ_0, and precision was estimated from 2α observations with sample mean μ_0 and sum of squared deviations 2β	$t_{2\alpha'}\left(\tilde{x}\mid\mu', \frac{\beta'(\nu'+1)}{\alpha'\nu'}\right)$[5]
Multivariate normal with known covariance matrix $\boldsymbol\Sigma$	$\boldsymbol\mu$ (mean vector)	Multivariate normal	$\boldsymbol\mu_0,\, \boldsymbol\Sigma_0$	$\left(\boldsymbol\Sigma_0^{-1} + n\boldsymbol\Sigma^{-1}\right)^{-1}\left(\boldsymbol\Sigma_0^{-1}\boldsymbol\mu_0 + n\boldsymbol\Sigma^{-1}\bar{\mathbf{x}}\right),$ $\left(\boldsymbol\Sigma_0^{-1} + n\boldsymbol\Sigma^{-1}\right)^{-1}$ • $\bar{\mathbf{x}}$ is the sample mean	mean was estimated from observations with total precision (sum of all individual precisions) $\boldsymbol\Sigma_0^{-1}$ and with sample mean $\boldsymbol\mu_0$	$\mathcal{N}(\tilde{\mathbf{x}}\mid{\boldsymbol\mu_0}', {\boldsymbol\Sigma_0}' + \boldsymbol\Sigma)$[5]
Multivariate normal with known precision matrix $\boldsymbol\Lambda$	$\boldsymbol\mu$ (mean vector)	Multivariate normal	$\boldsymbol\mu_0,\, \boldsymbol\Lambda_0$	$\left(\boldsymbol\Lambda_0 + n\boldsymbol\Lambda\right)^{-1}\left(\boldsymbol\Lambda_0\boldsymbol\mu_0 + n\boldsymbol\Lambda\bar{\mathbf{x}}\right),\, \left(\boldsymbol\Lambda_0 + n\boldsymbol\Lambda\right)$ • $\bar{\mathbf{x}}$ is the sample mean	mean was estimated from observations with total precision (sum of all individual precisions) $\boldsymbol\Lambda_0$ and with sample mean $\boldsymbol\mu_0$	$\mathcal{N}\left(\tilde{\mathbf{x}}\mid{\boldsymbol\mu_0}', ({\boldsymbol\Lambda_0}'^{-1} + \boldsymbol\Lambda^{-1})^{-1}\right)$[5]
Multivariate normal with known mean $\boldsymbol\mu$	$\boldsymbol\Sigma$ (covariance matrix)	Inverse-Wishart	$\nu,\, \boldsymbol\Psi$	$n+\nu,\, \boldsymbol\Psi + \sum_{i=1}^n (\mathbf{x_i}-\boldsymbol\mu)(\mathbf{x_i}-\boldsymbol\mu)^T$	covariance matrix was estimated from ν observations with sum of pairwise deviation products $\boldsymbol\Psi$	$t_{\nu'-p+1}\left(\tilde{\mathbf{x}}\vert\boldsymbol\mu, \frac{1}{\nu'-p+1}\boldsymbol\Psi'\right)$[5]
Multivariate normal with known mean $\boldsymbol\mu$	$\boldsymbol\Lambda$ (precision matrix)	Wishart	$\nu,\, \mathbf{V}$	$n+\nu,\, \left(\mathbf{V}^{-1} + \sum_{i=1}^n (\mathbf{x_i}-\boldsymbol\mu)(\mathbf{x_i}-\boldsymbol\mu)^T\right)^{-1}$	covariance matrix was estimated from ν observations with sum of pairwise deviation products $\mathbf{V}^{-1}$	$t_{\nu'-p+1}\left(\tilde{\mathbf{x}}\mid\boldsymbol\mu, \frac{1}{\nu'-p+1}{\mathbf{V}'}^{-1}\right)$[5]