Xinyu Tan

- **1. Exchangeability with known model parameters**

  (a) $p(y_1 = \text{b}, y_2 = \text{w}) = p(y_1 = \text{w}, y_2 = \text{b}) = 1/4$. Observations y_1 and y_2 are exchangable. We put the balls back, ergo y_1 and y_2 are independent.

  (b) y_1 and y_2 are exchangable, but not independent.

  (c) y_1 and y_2 are exchangable, but not independent. However, there are 1 million balls for each color, and we only take twice. Hence, we can act as if two observations are independent.

- **5. Mixtures of independent distributions**

  W.L.O.G, let's compute the covariance of θ_i and θ_j:

$$
\begin{aligned}
\text{cov}(\theta_i, \theta_j) &= \int_{\theta_i}\int_{\theta_j} p(\theta_i, \theta_j)(\theta_i - \mu_i)(\theta_j - \mu_j) d\theta_i d\theta_j \\
&= \int_{\alpha}\int_{\theta_i}\int_{\theta_j} p(\theta_i, \theta_j|\alpha)p(\alpha)(\theta_i - \mu_i)(\theta_j - \mu_j) d\alpha d\theta_i d\theta_j \\
&= \int_{\alpha}\int_{\theta_i}\int_{\theta_j} p(\theta_i|\alpha)(\theta_i - \mu_i) d\theta_i p(\theta_j|\alpha)(\theta_j - \mu_j) d\theta_j p(\alpha) d\alpha \\
&= \int_{\alpha}\left(\int_{\theta} p(\theta|\alpha)(\theta - \mu) d\theta\right)^2 p(\alpha) d\alpha \\
&\geq 0
\end{aligned}
$$

- **7. Continuous mixture models**

  (a) We have

$$
\begin{aligned}
y|\theta &\sim \frac{1}{y!}\theta^y e^{-\theta} \\
\theta|\alpha, \beta &\sim \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta}
\end{aligned}
$$

  Therefore, the prior predictive distribution of y:

$$
\begin{aligned}
p(y|\alpha, \beta) &= \int_{\theta} p(y, \theta|\alpha, \beta) d\theta \\
&= \int_{\theta} p(y|\theta)p(\theta|\alpha, \beta) d\theta \\
&= \frac{\beta^\alpha}{y!\Gamma(\alpha)}\int \theta^{\alpha+y-1}e^{-(\beta+1)\theta} d\theta \\
&= \frac{\Gamma(\alpha + y)}{y!\Gamma(\alpha)}\frac{\beta^\alpha}{(\beta+1)^{\alpha+y}}\int \frac{(\beta+1)^{\alpha+y}}{\Gamma(\alpha+y)}\theta^{\alpha+y-1}e^{-(\beta+1)\theta} d\theta \\
&= \frac{\Gamma(\alpha + y)}{y!\Gamma(\alpha)}\left(\frac{\beta}{\beta+1}\right)^\alpha\left(\frac{1}{\beta+1}\right)^y
\end{aligned}
$$

  Hence, the negative binomial.

To calculate the mean and variance of $y|\alpha,\beta$, we have

$$\mathbb{E}(y) = \mathbb{E}\left(\mathbb{E}(y|\theta)\right) = \mathbb{E}(\theta) = \frac{\alpha}{\beta}$$

and

$$\mathbb{V}(y) = \mathbb{E}\left(\mathbb{V}(y|\theta)\right) + \mathbb{V}\left(\mathbb{E}(y|\theta)\right) = \mathbb{E}(\theta) + \mathbb{V}(\theta) = \frac{\alpha}{\beta^2}(\beta+1)$$

(b) From Eq.(3.3), we have $\mu|\sigma^2, y \sim N(\bar{y}, \sigma^2/n)$. Therefore,

$$\mathbb{E}(\sqrt{n}(\mu-\bar{y})/s) = \mathbb{E}\left(\mathbb{E}(\sqrt{n}(\mu-\bar{y})/s|\sigma^2)\right) = \mathbb{E}\left(\sqrt{n}\mathbb{E}(\mu-\bar{y}|\sigma^2,y)/s^2\right) = 0$$

and

$$\begin{aligned}
\mathbb{V}(\sqrt{n}(\mu-\bar{y})/s) &= \mathbb{E}\left(\mathbb{V}(\sqrt{n}(\mu-\bar{y})/s|\sigma^2)\right) + \mathbb{V}\left(\mathbb{E}(\sqrt{n}(\mu-\bar{y})/s|\sigma^2)\right) \\
&= \mathbb{E}\left(\frac{n}{s^2}\mathbb{V}(\mu|\sigma^2,y)\right) + 0 \\
&= \mathbb{E}\left(\frac{n}{s^2}\frac{\sigma^2}{n}\right) \qquad \sigma^2|y \sim \text{Inv-}\chi^2(n-1,s^2) \\
&= \frac{n-1}{n-3}
\end{aligned}$$

– **9. Noninformative hyperprior distribution**

(a)

(b) We have uniform distribution on $\left(\frac{\alpha}{\alpha+\beta}, (\alpha+\beta)^{-1/2}\right)$. Denote $u = \frac{\alpha}{\alpha+\beta}$ and $v = (\alpha+\beta)^{-1/2}$, then

$$p(\alpha,\beta) = p(u,v)\begin{vmatrix} \frac{\partial u}{\partial \alpha} & \frac{\partial u}{\partial \beta} \\ \frac{\partial v}{\partial \alpha} & \frac{\partial v}{\partial \beta} \end{vmatrix} = \begin{vmatrix} \beta/(\alpha+\beta)^2 & -\alpha/(\alpha+\beta)^2 \\ -\frac{1}{2}(\alpha+\beta)^{-3/2} & -\frac{1}{2}(\alpha+\beta)^{-3/2} \end{vmatrix} = -\frac{1}{2}(\alpha+\beta)^{-5/2}$$