Bayesian Statistics

Tutorial 6

YANG Chen[1] NG Say-Yao Mark[2] CHOO Kuan Li Justin[3]

National University of Singapore (NUS)

October 7, 2016

A Short Introduction

A Short Introduction

- ☺ NG Say-Yao Mark: Year 4, double major in Mathematics and Statistics, minor in Computer Science.

A Short Introduction

- ☺ NG Say-Yao Mark: Year 4, double major in Mathematics and Statistics, minor in Computer Science.
- ☺ YANG Chen: Year 4, major in Statistics.

A Short Introduction

- ☺ NG Say-Yao Mark: Year 4, double major in Mathematics and Statistics, minor in Computer Science.
- ☺ YANG Chen: Year 4, major in Statistics.
- ☺ CHOO Kuan Li Justin: Year 3, major in Statistics, minor in Economics.

Question 1 (a)

Suppose that we observe two independent normal samples,

$$(x_1, \ldots, x_m) \overset{\text{IID}}{\sim} \mathcal{N}(\mu_1, \sigma_1^2) \text{ and } (y_1, \ldots, y_n) \overset{\text{IID}}{\sim} \mathcal{N}(\mu_2, \sigma_2^2).$$

Suppose also that the parameters $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2)$ are assigned the vague prior

$$p(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \propto \frac{1}{\sigma_1^2 \sigma_2^2}.$$

Find the joint posterior density of $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2)$ and show that (μ_1, σ_1^2) and (μ_2, σ_2^2) have independent posterior distributions. Describe how to simulate from this joint posterior density.

Question 1 (a): Solution

First note that the individual priors $p(\mu_1, \sigma_1^2)$ and $p(\mu_2, \sigma_2^2)$ are independent since

$$p(\mu_1, \sigma_1^2) \propto \frac{1}{\sigma_1^2} \text{ and } p(\mu_2, \sigma_2^2) \propto \frac{1}{\sigma_2^2}.$$

The joint likelihood of $\mathbf{x} = (x_1, \ldots, x_m)$ and $\mathbf{y} = (y_1, \ldots, y_n)$ can be decomposed into

$$p(\mathbf{x}, \mathbf{y} | \mu_1, \sigma_1^2, \mu_2, \sigma_2^2) = p(\mathbf{x} | \mu_1, \sigma_1^2) p(\mathbf{y} | \mu_2, \sigma_2^2).$$

Question 1 (a): Solution

Therefore

$$\begin{aligned} p(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2 | \mathbf{x}, \mathbf{y}) &\propto p(\mathbf{x}, \mathbf{y} | \mu_1, \sigma_1^2, \mu_2, \sigma_2^2) p(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \\ &\propto p(\mathbf{x} | \mu_1, \sigma_1^2) p(\mu_1, \sigma_1^2) p(\mathbf{y} | \mu_2, \sigma_2^2) p(\mu_2, \sigma_2^2). \end{aligned}$$

This implies that the individual posterior distributions are independent since

$$p(\mu_1, \sigma_1^2 | \mathbf{x}) \propto p(\mathbf{x} | \mu_1, \sigma_1^2) p(\mu_1, \sigma_1^2),$$

and

$$p(\mu_2, \sigma_2^2 | \mathbf{y}) \propto p(\mathbf{y} | \mu_2, \sigma_2^2) p(\mu_2, \sigma_2^2).$$

Question 1 (a): Solution

So we consider the individual posterior distributions.

Consider $\mathbf{x}, \mu_1, \sigma_1^2$ first. Note that

$$\sum_{i=1}^{m} \frac{(x_i - \mu_1)^2}{2\sigma_1^2} = \frac{1}{2\sigma_1^2} \left(m(\mu_1 - \bar{x})^2 + \sum_{i=1}^{m} (x_i - \bar{x})^2 \right).$$

Question 1 (a): Solution

This is implies that:

$$
\begin{aligned}
p(\mu_1, \sigma_1^2 | \mathbf{x}) &\propto p(\mathbf{x} | \mu_1, \sigma_1^2) p(\mu_1, \sigma_1^2) \\
&= \left(\prod_{i=1}^{m} \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left\{ -\frac{(x_i - \mu_1)^2}{2\sigma_1^2} \right\} \right) \cdot \frac{1}{\sigma_1^2} \\
&\propto \frac{1}{\sigma_1^{m+2}} \exp\left\{ -\sum_{i=1}^{m} \frac{(x_i - \mu_1)^2}{2\sigma_1^2} \right\} \\
&\propto \frac{1}{(\sigma_1^2)^{m/2+1}} \exp\left\{ -\frac{1}{2\sigma_1^2} \left(m(\mu_1 - \bar{x})^2 + \sum_{i=1}^{m} (x_i - \bar{x})^2 \right) \right\}
\end{aligned}
$$

Question 1 (a): Solution

(*continued from previous slide...*)

$$
\begin{aligned}
&\propto (\sigma_1^2)^{-m/2-1} \exp\left\{-\frac{\sum_{i=1}^m (x_i - \bar{x})^2}{2\sigma_1^2}\right\} \exp\left\{-\frac{m(\mu_1 - \bar{x})^2}{2\sigma_1^2}\right\} \\
&\propto p(\sigma_1^2 | \mathbf{x}) \cdot p(\mu_1 | \sigma_1^2, \mathbf{x}),
\end{aligned}
$$

where

$$
\begin{cases}
\sigma_1^2 | \mathbf{x} \sim \text{InvGamma}\left(m/2, \frac{\sum_{i=1}^m (x_i - \bar{x})^2}{2}\right) \\
\mu_1 | \sigma_1^2, \mathbf{x} \sim \mathcal{N}(\bar{x}, \sigma_1^2/m).
\end{cases}
$$

Question 1 (a): Solution

In a similar fashion we obtain:

$$\begin{cases} \sigma_2^2|\mathbf{y} \sim \text{InvGamma}\left(n/2, \frac{\sum_{j=1}^n (y_i - \bar{y})^2}{2}\right) \\ \mu_2|\sigma_2^2, \mathbf{y} \sim \mathcal{N}(\bar{y}, \sigma_2^2/n). \end{cases}$$

Question 1 (a): Solution

Since the posterior distribution of the pairs (μ_1, σ_1^2) and (μ_2, σ_2^2) are independent we generate the pairs individually.

- To generate (μ_1, σ_1^2):
  1. Generate $\sigma_1^2|\mathbf{x} \sim \text{InvGamma}\left(m/2, \frac{\sum_{i=1}^m (x_i - \bar{x})^2}{2}\right)$.
  2. Generate $\mu_1|\sigma_1^2, \mathbf{x} \sim \mathcal{N}(\bar{x}, \sigma_1^2/m)$.
- To generate (μ_2, σ_2^2), apply a similar procedure as above.
  1. Generate $\sigma_2^2|\mathbf{y} \sim \text{InvGamma}\left(n/2, \frac{\sum_{j=1}^n (y_i - \bar{y})^2}{2}\right)$.
  2. Generate $\mu_2|\sigma_2^2, \mathbf{y} \sim \mathcal{N}(\bar{y}, \sigma_2^2/n)$.

Question 1 (b)

The following data gives the mandible length in millimetres for 10 male and 10 female golden jackals in the British Museum.

Males										Females									
120	107	110	116	114	111	113	117	114	112	110	111	107	108	110	105	107	106	111	111

- Using the the model in 1(a), simulate 10,000 samples of $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2)$ from the joint posterior density.
- Construct a kernel estimate of the posterior density of the difference in mean mandible length between the sexes, using the `density` command in `R`.
- Check if there is sufficient evidence to conclude that males have a larger average length.

Question 1 (b): Solution

- Let $\mathbf{x}, \mu_1, \sigma_1^2$ correspond to the $m = 10000$ males.
- Let $\mathbf{y}, \mu_2, \sigma_2^2$ correspond to the $n = 10000$ females.

For the males:

```
> males <- c(120,107,110,116,114,111,113,117,114,112)
> SSE.male <- sum((males-mean(males))^2)
> sigma.1.squared <- 1/rgamma(n=10000,
+                              shape=10/2,
+                              rate=SSE.male/2)
> mu.1 <- rnorm(n=10000,
+               mean=mean(males),
+               sd=sqrt(sigma.1.squared/10))
```

Question 1 (b): Solution

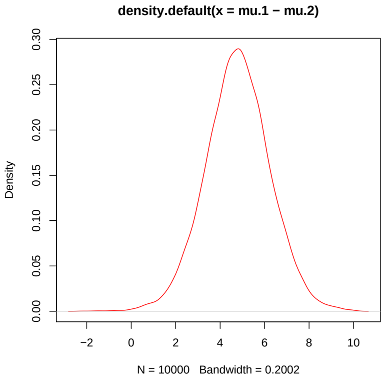
For the females:

```
> females <- c(110,111,107,108,110,105,107,106,111,111)
> SSE.female <- sum((females-mean(females))^2)
> sigma.2.squared <- 1/rgamma(n=10000,
+                             shape=10/2,
+                             rate=SSE.female/2)
> mu.2 <- rnorm(n=10000,
+               mean=mean(females),
+               sd=sqrt(sigma.2.squared/10))
```

Question 1 (b): Solution

To estimate the posterior density of the difference in mean length between the sexes, $\mu_1 - \mu_2$, perform the following:

```
> plot(density(mu.1 - mu.2), type='l', col='red')
```



Question 1 (b): Solution

To decide whether there is sufficient evidence to conclude that males have a larger average, we can setup the following hypotheses:

$$\begin{cases} H_0: & \mu_1 - \mu_2 > 0, \\ H_1: & \mu_1 - \mu_2 \leq 0. \end{cases}$$

We then compare the posterior odds, p_0/p_1 where

$$\begin{cases} & p_0 = \mathbf{P}(\mu_1 - \mu_2 > 0), \\ & p_1 = 1 - p_0. \end{cases}$$

Question 1 (b): Solution

The computation is shown below:

```
> p.0 <- sum(mu.1 - mu.2 > 0) / length(mu.1 - mu.2)
> p.1 <- 1 - p.0
> cat('Posterior odds:', p.0/p.1)
```

```
Posterior odds: 665.6667
```

The posterior odds are high, hence there is strong evidence supporting the hypothesis H_0 that males have a larger average.

Question 2

After a posterior analysis on data from a population of squash plants, it was determined that the weight of a randomly chosen play, Y, could be modelled with the distribution:

$$p(y, \theta, \sigma^2) = 0.31\phi(y|\theta, \sigma) + 0.46\phi(y|2\theta, 2\sigma) + 0.23\phi(y|3\theta, 3\sigma),$$

where $\phi(y|\theta, \sigma)$ denotes the PDF of a $\mathcal{N}(\theta, \sigma^2)$ distribution. Ther posterior distributions of the parameters have been calculated as

$$\sigma^2 \sim \text{InvGamma}(10, 2.5) \text{ and } \theta|\sigma^2 \sim \mathcal{N}(4.1, \sigma^2/20).$$

Question 2: Sampling from the Posterior Predictive Distribution

How to sample $10,000$ y-values from the posterior predictive distribution?

Note that the posterior predictive distribution is

$$p(\tilde{y}|y) = \int p(\tilde{y}|\theta,\sigma^2)p(\theta,\sigma^2|y)\,\mathrm{d}\theta\mathrm{d}\sigma^2 = \mathbf{E}_{\theta,\sigma^2|y}\left[p(\tilde{y}|\theta,\sigma^2)\right],$$

where

$$p(\tilde{y},\theta,\sigma^2) = 0.31\phi(\tilde{y}|\theta,\sigma) + 0.46\phi(\tilde{y}|2\theta,2\sigma) + 0.23\phi(\tilde{y}|3\theta,3\sigma).$$

Question 2: Sampling from the Posterior Predictive Distribution

Our proposed steps to sampling $p(\tilde{y}|y)$ is:

(1) Sample $\sigma^2 \sim \text{InvGamma}(10, 2.5)$.

(2) Sample $\theta|\sigma^2 \sim \mathcal{N}(4.1, \sigma^2/20)$.

(3) Sample $\tilde{y}|\theta, \sigma^2$ based on

$$p(\tilde{y}|\theta, \sigma^2) = 0.31\phi(\tilde{y}|\theta, \sigma) + 0.46\phi(\tilde{y}|2\theta, 2\sigma) + 0.23\phi(\tilde{y}|3\theta, 3\sigma).$$

- ⋆ Repeat steps (1)-(3) until a large enough sample of $\tilde{y}$ is generated.
- ⋆ Use this sample as an estimate of the posterior predicitve distribution.

Question 2: Sampling from the Posterior Predictive Distribution

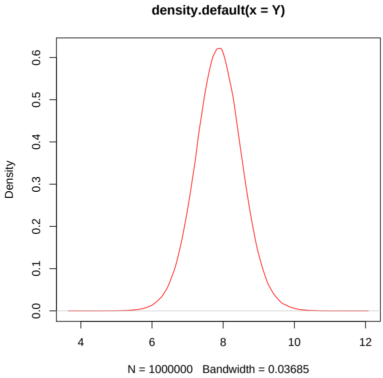
The computation in R is shown below:

```
> sigma.squared <- 1/rgamma(n=1e6, shape=10, rate=2.5)
> sigma <- sqrt(sigma.squared)
> theta <- rnorm(n=1e6, mean=4.1,
+                     sd=sqrt(sigma.squared/20))
> Y <- (   .31*rnorm(n=1e6, mean=theta, sd=sigma)
+        + .46*rnorm(n=1e6, mean=2*theta, sd=2*sigma)
+        + .23*rnorm(n=1e6, mean=3*theta, sd=3*sigma) )
```

Question 2: Sampling from the Posterior Predictive Distribution

The estimate of $p(\tilde{y}|y)$ is shown below:

```
> plot(density(Y), type='l', col='red')
```



Question 2 (a) & Solution

(a) Find a 75% quantile-based confidence interval for a new value of Y.

Solution. We use the sample distribution generated previously.

```
> cat('75% Quantile-based CI:',
+       '[',
+       quantile(Y, .125),
+       ',',
+       quantile(Y, .875),
+       ']')
```

```
75% Quantile-based CI: [ 7.122648 , 8.622392 ]
```

Question 2 (b)

Find a 75% HPD region for a new Y as follows:

(i) Compute estimates of the posterior density of Y using the `density` command in `R` and then normalize the density values so that the sum to 1. Sort these discrete probabilites in decreasing order.

(ii) Find the first probability value such that the cumulative sum of the sorted values exceeds 0.75. Your HPD region includes all values of y which have a discretized probability greater than this cutoff. Describe your HPD region and compare it to your quantile-based region.

Question 2 (b): Solutions

For (i), we carry out the following computation in R.

```
> values <- density(Y)$y / sum(density(Y)$y)
> values.sorted <- sort(values, decreasing=TRUE)
```

Question 2 (b): Solutions

For (ii), we carry out the following computation in R.

```
> cumulative.sum <- cumsum(values.sorted)
> index <- 1
> while (cumulative.sum[index] < 0.75) {
+   index <- index + 1
+ }
> probability.value <- values.sorted[index]
> HPD <- density(Y)$x[values > probability.value]
> cat('75% HPD:', '[', min(HPD), ',', max(HPD), ']')
```

```
75% HPD: [ 7.13184 , 8.616311 ]
```

The 75% HPD is similar to the 75% quantile-based CI.

Question 3

Assume that $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ is an IID sample generated from $\mathbf{x} = (x_1, x_2, x_3) \sim \text{Multinomial}(1; p_1, p_2, p_3)$ which has PDF

$$p(\mathbf{x}|p_1, p_2, p_3) = \frac{1!}{x_1!x_2!x_3!} p_1^{x_1} p_2^{x_2} p_3^{x_3}.$$

with $x_j \in [0, 1]$ for $j = 1, 2, 3$, $x_1 + x_2 + x_3 = 1$ and $p_j \geq 0$ for $j = 1, 2, 3$, $p_1 = p_2 + p_3 = 1$. Assume a Dirichlet prior with nonnegative parameters α_j, $j = 1, 2, 3$ for (p_1, p_2, p_3):

$$p(p_1, p_2, p_3) = \frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_2)} p_1^{\alpha_1 - 1} p_2^{\alpha_2 - 1} p_3^{\alpha_3 - 1}.$$

Let $\mathbf{x}_i = (x_{i1}, x_{i2}, x_{i3})$, $i = 1, 2, \ldots, n$ and set $x_{\cdot j} = \sum_{i=1}^{n} x_{ij}$, $j = 1, 2, 3$.

Question 3

(a) For the posterior PDF $p(p_1, p_2, p_3|\mathbf{x}_1, \ldots, \mathbf{x}_n)$ and show that the prior $p(p_1, p_2, p_3)$ is a conjugate prior for (p_1, p_2, p_3) with respect to $p(\mathbf{x}|p_1, p_2, p_3)$.

(b) Let $\mathbf{x}_{n+1}$ be a new observation. Derive the PDF $p(\mathbf{x}_{n+1}|\mathbf{x}_1, \ldots, \mathbf{x}_n)$ and identify the distribution. Note that $\Gamma(a+1) = a\Gamma(a)$ for $a > 0$ and $1! = 0!$.

Question 3 (a): Solution

First we derive the joint likelihood of $\mathbf{x}_1, \ldots, \mathbf{x}_n$:

$$
\begin{aligned}
& p(\mathbf{x}_1, \ldots, \mathbf{x}_n | p_1, p_2, p_3) \\
&= \prod_{i=1}^{n} \left(\frac{1!}{x_{i1}!x_{i2}!x_{i3}!} p_1^{x_{i1}} p_2^{x_{i2}} p_3^{x_{i3}} \right) \\
&= \left(\prod_{i=1}^{n} \frac{1!}{x_{i1}!x_{i2}!x_{i3}!} \right) p_1^{\sum_{i=1}^{n} x_{i1}} p_2^{\sum_{i=1}^{n} x_{i2}} p_3^{\sum_{i=1}^{n} x_{i3}}.
\end{aligned}
$$

Question 3 (a): Solution

Recall that

$$p(p_1, p_2, p_3) = \frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_3}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)} p_1^{\alpha_1 - 1} p_2^{\alpha_2 - 1} p_3^{\alpha_3 - 1}.$$

Question 3 (a): Solution

Therefore

$$
\begin{aligned}
& p(p_1, p_2, p_3 | \mathbf{x}_1, \ldots, \mathbf{x}_n) \\
\propto \quad & p(\mathbf{x}_1, \ldots, \mathbf{x}_n | p_1, p_2, p_3) p(p_1, p_2, p_3) \\
\propto \quad & p_1^{\sum_{i=1}^n x_{i1}} p_2^{\sum_{i=1}^n x_{i2}} p_3^{\sum_{i=1}^n x_{i3}} \cdot p_1^{\alpha_1 - 1} p_2^{\alpha_2 - 1} p_3^{\alpha_3 - 1} \\
\propto \quad & p_1^{\alpha_1 - 1 + \sum_{i=1}^n x_{i1}} p_2^{\alpha_2 - 1 + \sum_{i=1}^n x_{i2}} p_3^{\alpha_3 - 1 + \sum_{i=1}^n x_{i3}}
\end{aligned}
$$

From this we can identify the posterior density to be Dirichlet with parameters $\alpha_j + \sum_{i=1}^n x_{ij}$ for $j = 1, 2, 3$. So it is clear that our prior is a conjugate prior for (p_1, p_2, p_3) with respect to $p(\mathbf{x} | p_1, p_2, p_3)$

Question 3 (b): Solution

By definition of posterior predictive distribution we know that $p(\mathbf{x}_{n+1}|\mathbf{x}_1, \ldots, \mathbf{x}_n)$ is equal to

$$\int p(\mathbf{x}_{n+1}|p_1, p_2, p_3)p(p_1, p_2, p_3|\mathbf{x}_1, \ldots, \mathbf{x}_n)\, \mathrm{d}p_1 \mathrm{d}p_2 \mathrm{d}p_3,$$

where the integral ranges over the prescribed domain of (p_1, p_2, p_3), i.e. all $p_1, p_2, p_3 \in [0, 1]$ such that $p_1 + p_2 + p_3 = 1$.

Question 3 (b): Solution

Let

$$C_\alpha = \frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_3 + \sum_{i=1}^n x_{i1} + x_{i2} + x_{i3})}{\Gamma(\alpha_1 + \sum_{i=1}^n x_{i1})\Gamma(\alpha_2 + \sum_{i=1}^n x_{i2})\Gamma(\alpha_3 + \sum_{i=1}^n x_{i3})}$$
$$= \frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_3 + n)}{\Gamma(\alpha_1 + \sum_{i=1}^n x_{i1})\Gamma(\alpha_2 + \sum_{i=1}^n x_{i2})\Gamma(\alpha_3 + \sum_{i=1}^n x_{i3})}$$

and

$$C_x = \frac{1!}{x_{n+1,1}!x_{n+1,2}!x_{n+1,3}!} = 1.$$

Question 3 (b): Solution

Then the integral reduces to:

$$C_\alpha C_x \int p_1^{\alpha_1 - 1 + \sum_{i=1}^{n+1} x_{i,1}} p_2^{\alpha_2 - 1 + \sum_{i=1}^{n+1} x_{i,2}} p_3^{\alpha_3 - 1 + \sum_{i=1}^{n+1} x_{i3}} \mathrm{d}p_1 \mathrm{d}p_2 \mathrm{d}p_3.$$

Question 3 (b): Solution

We evaluate the integral, using the fact that the Dirichlet density integrates to one; so it must give the appropriate normalization constant.

$$
\begin{aligned}
& C_\alpha C_x \int p_1^{\alpha_1 - 1 + \sum_{i=1}^{n+1} x_{i,1}} p_2^{\alpha_2 - 1 + \sum_{i=1}^{n+1} x_{i,2}} p_3^{\alpha_3 - 1 + \sum_{i=1}^{n+1} x_{i3}} \mathrm{d}p_1 \mathrm{d}p_2 \mathrm{d}p_3 \\
&= C_\alpha \frac{\Gamma(\alpha_1 + \sum_{i=1}^{n+1} x_{i1})\Gamma(\alpha_2 + \sum_{i=1}^{n+1} x_{i2})\Gamma(\alpha_3 + \sum_{i=1}^{n+1} x_{i3})}{\Gamma(\alpha_1 + \alpha_2 + \alpha_3 + \sum_{i=1}^{n+1} x_{i1} + x_{i2} + x_{i3})} \\
&= C_\alpha \frac{\Gamma(\alpha_1 + \sum_{i=1}^{n+1} x_{i1})\Gamma(\alpha_2 + \sum_{i=1}^{n+1} x_{i2})\Gamma(\alpha_3 + \sum_{i=1}^{n+1} x_{i3})}{\Gamma(\alpha_1 + \alpha_2 + \alpha_3 + n + 1)} \\
&= C_\alpha \frac{\Gamma(\alpha_1 + \sum_{i=1}^{n+1} x_{i1})\Gamma(\alpha_2 + \sum_{i=1}^{n+1} x_{i2})\Gamma(\alpha_3 + \sum_{i=1}^{n+1} x_{i3})}{(\alpha_1 + \alpha_2 + \alpha_3 + n)\Gamma(\alpha_1 + \alpha_2 + \alpha_3 + n)}
\end{aligned}
$$

Question 3 (b): Solution

(*continued from previous slide*)
Notice that only one of the variables x_{i1}, x_{i2}, x_{i3} can be equal to 1. Whereas the remaining two ust be set to 0.

$\implies$ Without loss of generality we may assume $x_{i1} = 1$ whilst $x_{i2} = x_{i3} = 0$. Then we get the following cancellation of terms,

$$
\begin{aligned}
& C_\alpha \frac{\Gamma(\alpha_1 + \sum_{i=1}^{n+1} x_{i1})\Gamma(\alpha_2 + \sum_{i=1}^{n+1} x_{i2})\Gamma(\alpha_3 + \sum_{i=1}^{n+1} x_{i3})}{(\alpha_1 + \alpha_2 + \alpha_3 + n)\Gamma(\alpha_1 + \alpha_2 + \alpha_3 + n)} \\
=& C_\alpha \frac{\Gamma(\alpha_1 + \sum_{i=1}^{n+1} x_{i1})\Gamma(\alpha_2 + \sum_{i=1}^{n} x_{i2})\Gamma(\alpha_3 + \sum_{i=1}^{n} x_{i3})}{(\alpha_1 + \alpha_2 + \alpha_3 + n)\Gamma(\alpha_1 + \alpha_2 + \alpha_3 + n)} \\
=& \frac{\Gamma(\alpha_1 + \sum_{i=1}^{n+1} x_{i1})}{\Gamma(\alpha_1 + \sum_{i=1}^{n} x_{i1})(\alpha_1 + \alpha_2 + \alpha_3 + n)}
\end{aligned}
$$

Question 3 (b): Solution

(*continued from previous slide*)

$$
\begin{aligned}
&\vdots \\
&= \frac{\Gamma(\alpha_1 + \sum_{i=1}^{n+1} x_{i1})}{\Gamma(\alpha_1 + \sum_{i=1}^{n} x_{i1})(\alpha_1 + \alpha_2 + \alpha_3 + n)} \\
&= \frac{(\alpha_1 + \sum_{i=1}^{n} x_{i1})\Gamma(\alpha_1 + \sum_{i=1}^{n} x_{i1})}{\Gamma(\alpha_1 + \sum_{i=1}^{n} x_{i1})(\alpha_1 + \alpha_2 + \alpha_3 + n)} \\
&= \frac{(\alpha_1 + \sum_{i=1}^{n} x_{i1})}{(\alpha_1 + \alpha_2 + \alpha_3 + n)}
\end{aligned}
$$

Question 3 (b): Solution

This means that $\mathbf{x}_{n+1} \in \{(1,0,0),(0,1,0),(0,0,1)\}$ follows the multinomial distribution with

$$\begin{cases} \mathbf{P}(\mathbf{x}_{n+1} = (1,0,0)) = \frac{(\alpha_1 + \sum_{i=1}^{n} x_{i1})}{(\alpha_1+\alpha_2+\alpha_3+n)} \\ \mathbf{P}(\mathbf{x}_{n+1} = (0,1,0)) = \frac{(\alpha_2 + \sum_{i=1}^{n} x_{i2})}{(\alpha_1+\alpha_2+\alpha_3+n)} \\ \mathbf{P}(\mathbf{x}_{n+1} = (0,0,1)) = \frac{(\alpha_3 + \sum_{i=1}^{n} x_{i3})}{(\alpha_1+\alpha_2+\alpha_3+n)} \end{cases}$$

$\square$

Thank you very much for your attention.