$Y_1, ..., Y_n$ are *exchangeable*
$\Rightarrow$
$Y_1, ..., Y_n|\theta$ are conditionally i.i.d., $\theta \sim \pi(\theta)$ (de Finetti)
$\Rightarrow$

$$p(\tilde{Y} = \tilde{y}|Y_1 = y_1, ..., Y_n = y_n) = \frac{p(\tilde{y}, y_1, ..., y_n)}{p(y_1, ..., y_n)} \tag{1}$$

$$= \frac{\int p(\tilde{y}, y_1, ..., y_n|\theta)\pi(\theta)d\theta}{\int p(y_1, ..., y_n|\theta)\pi(\theta)d\theta} \tag{2}$$

$$= \frac{\int p(\tilde{y}|\theta)p(y_1, ..., y_n|\theta)\pi(\theta)d\theta}{\int p(y_1, ..., y_n|\theta)\pi(\theta)d\theta} \tag{3}$$

$$= \frac{\int p(\tilde{y}|\theta)p(\theta|y_1, ..., y_n)p(y_1, ..., y_n)d\theta}{\int p(y_1, ..., y_n|\theta)\pi(\theta)d\theta} \tag{4}$$

$$= \frac{\cancel{p(y_1, ..., y_n)} \int p(\tilde{y}|\theta)p(\theta|y_1, ..., y_n)d\theta}{\cancel{p(y_1, ..., y_n)}} \tag{5}$$

$$= \int p(\tilde{y}|\theta)p(\theta|y_1, ..., y_n)d\theta \tag{6}$$

Explanation:

1. The conditional density of $\tilde{y}$ is the quotient of the joint density $\tilde{y}, y_1, ..., y_n$ with the marginal joint density of $y_1, ..., y_n$ (definition of conditional density)

2. The joint densities in both the numerator and denominator can be thought of as marginal joint densities, and written as integrals (with respect to something–here we conveniently choose θ) of products of conditional joint densities (conditional on θ) with the density of θ

3. Conditional independence of the y_is with respect to θ enables us to express the conditional joint density in the numerator as a product of densities

4. Pairing the second and third factor in the integrand of the numerator, we apply Bayes' rule ($p(A|B)p(B) = p(B|A)p(A)$)

5. Now the last factor in the integrand of the numerator does not depend on θ, and can be pulled out of the integral, while the denominator just reverts back to the original joint density from line 1 (clearly it was not necessary to change it in the first place), and these two quantities conveniently cancel out

6. We now have the Bayesian Predictive Inference format we're aiming for

---

[Showing here that it is indeed a NB distribution]

$$\theta \sim NB(\alpha, \beta) \Rightarrow p(\theta) = \binom{\theta + \alpha - 1}{\alpha - 1} \left(\frac{\beta}{\beta + 1}\right)^{\alpha} \left(\frac{1}{\beta + 1}\right)^{\theta}$$

so

$$\tilde{y} \sim NB\left(\alpha + \sum y_i), \beta + n\right) \Rightarrow p(\tilde{y}) = \binom{\tilde{y} + \alpha + \sum y_i - 1}{\alpha + \sum y_i - 1} \left(\frac{\beta + n}{\beta + n + 1}\right)^{\alpha + \sum y_i} \left(\frac{1}{\beta + n + 1}\right)^{\tilde{y}}$$

$$= \frac{(\alpha + \sum y_i + \tilde{y} - 1)!}{(\alpha + \sum y_i - 1)!\,(\tilde{y})!} \left(\frac{\beta + n}{\beta + n + 1}\right)^{\alpha + \sum y_i} \left(\frac{1}{\beta + n + 1}\right)^{\tilde{y}}$$

$$= \frac{\Gamma\left(\alpha + \sum y_i + \tilde{y}\right)}{\Gamma\left(\alpha + \sum y_i\right)\Gamma\left(\tilde{y} + 1\right)} \left(\frac{\beta + n}{\beta + n + 1}\right)^{\alpha + \sum y_i} \left(\frac{1}{\beta + n + 1}\right)^{\tilde{y}}$$

, for which

$$E\left[\tilde{Y}|y_1, ..., y_n\right] = \frac{a + \sum y_i}{b + n} = E\left[\theta|y_1, ..., y_n\right];$$

$$\mathrm{Var}\left[\tilde{Y}|y_1, ..., y_n\right] = \frac{a + \sum y_i}{b + n}\frac{b + n + 1}{b + n}$$

$$= \mathrm{Var}\left[\theta|y_1, ..., y_n\right] \times (b + n + 1)$$

$$= E\left[\theta|y_1, ..., y_n\right] \times \frac{b + n + 1}{b + n}$$

we obtain the posterior

$$p(\theta|Y_1,...,Y_n) = \frac{p(Y_1,...,Y_n|\theta)\pi(\theta)}{\int p(Y_1,...,Y_n|\theta)\pi(\theta)d\theta}$$

$$= \frac{\theta^d e^{-\theta n\bar{y}} \cdot \frac{\gamma^\delta \theta^{\delta-1} e^{-\gamma\theta}}{\Gamma(\delta)}}{\int \left(\theta^d e^{-\theta n\bar{y}} \cdot \frac{\gamma^\delta \theta^{\delta-1} e^{-\gamma\theta}}{\Gamma(\delta)}\right) d\theta}$$

$$= \frac{\cancel{\frac{\gamma^\delta}{\Gamma(\delta)}} \left(\theta^{d+\delta-1} e^{-\theta(\gamma+n\bar{y})}\right)}{\cancel{\frac{\gamma^\delta}{\Gamma(\delta)}} \int \left(\theta^{d+\delta-1} e^{-\theta(\gamma+n\bar{y})}\right) d\theta}$$

$$= \frac{\frac{(\gamma+n\bar{y})^{d+\delta}}{\Gamma(d+\delta)} \left(\theta^{d+\delta-1} e^{-\theta(\gamma+n\bar{y})}\right)}{\cancel{\frac{(\gamma+n\bar{y})^{d+\delta}}{\Gamma(d+\delta)} \int \left(\theta^{d+\delta-1} e^{-\theta(\gamma+n\bar{y})}\right) d\theta}}$$

$$= \frac{(\gamma+n\bar{y})^{d+\delta} \theta^{d+\delta-1} e^{-\theta(\gamma+n\bar{y})}}{\Gamma(d+\delta)}$$

with the Gamma$(d+\delta, \gamma+n\bar{y})$ density in the next to last step integrating to 1.