Winter 19 – AMS206B Homework 1 Solution

1. Let $X \sim \text{Exp}(\lambda)$, where $E(X) = 1/\lambda$. What is the pmf (probability mass function) of $Y = \lfloor X \rfloor$ (the floor of X)? Do you recognize it as a distribution that you have studied in the past?

We first find $\Pr(Y = y) = \Pr(y \le X < y - 1) = \int_y^{y+1} \lambda e^{-\lambda x} dx = e^{-\lambda y} (1 - e^{-\lambda}), y = 0, 1, 2, \dots$ We observe it is the probability function for the geometric distribution with parameter $e^{-\lambda}$, i.e., $Y \sim \text{Geometric}(e^{-\lambda})$.

- 2. Let X_1 and X_2 be two independent random variables such that $X_i \sim \operatorname{Ga}(a_i, b)$ for any $a_1, a_2, b > 0$. Define $Y = X_1/(X_1 + X_2)$ and $Z = (X_1 + X_2)$.
 - (a) Find the joint pdf for Y and Z and show that these two random variables are independent.

 $Y = X_1/(X_1 + X_2)$ and $Z = (X_1 + X_2)$, $X_1, X_2 \in \mathbb{R}^+$ are differentiable one-to-one transformations. We observe $X_1 = YZ$ and $X_2 = Z(1 - Y)$. We use a change of variables. We have 0 < Y < 1 and $Z \in \mathbb{R}^+$, and find

$$\begin{split} g(y,z) &=& f(yz,z(1-y))|J| \\ &=& \frac{b^{a_1}}{\Gamma(a_1)}(yz)^{a_1-1}e^{-byz}\frac{b^{a_2}}{\Gamma(a_2)}((1-y)z)^{a_2-1}e^{-b(1-y)z}z \\ &\propto& \underbrace{y^{a_1-1}(1-y)^{a_2-1}}_{\text{function of }y\text{ only}}\underbrace{z^{a_1+a_2-1}e^{-bz}}_{\text{function of }z\text{ only}}, y \in (0,1) \text{ and } 0 < z. \end{split}$$

Since the joint pdf of Y and Z can be expressed as a product of a function of Y only and a function of Z only, Y and Z are independent.

(b) Find the marginal pdf of Z. Do you recognize this pdf as belonging to some family that you know?

From part (a), we have $f(z) \propto z^{a_1+a_2-1}e^{-bz}$, $z \in \mathbb{R}^+$ and recognize it is proportional to the pdf of $\operatorname{Gamma}(a_1+a_2,b)$. Thus, $Z \sim \operatorname{Gamma}(a_1+a_2,b)$.

(c) Find the marginal pdf of Y. Do you recognize this pdf as belonging to some family that you know?

From part (a), we have $f(y) \propto y^{a_1-1}(1-y)^{a_2-1}$, 0 < y < 1 and recognize it is proportional to the pdf of Be (a_1, a_2) . Thus, $Y \sim \text{Be}(a_1, a_2)$.

(d) Compute $E(Y^k)$ for any k > 0.

We have

$$E(Y^k) = \int_0^\infty y^k \frac{1}{B(a_1, a_2)} y^{a_1 - 1} (1 - y)^{a_2 - 1} dy = \frac{B(a_1 + k, a_2)}{B(a_1, a_2)}.$$

1

(e) What does this result imply if $a_i = b = 1$?

We observe $Y \sim \text{Be}(1,1) = \text{Unif}(1,1)$. For $Z = X_1 + \dots, X_n, Z \sim \text{Gamma}(n,1)$, that is, Erlang distribution.

3. Consider three independent random variables X_1 , X_2 and X_3 such that $X_i \overset{indep}{\sim} \text{Gamma}(a_i, b)$, i = 1, 2, 3. Let

$$Y = (Y_1, Y_2, Y_3) = \left(\frac{X_1}{X_1 + X_2 + X_3}, \frac{X_2}{X_1 + X_2 + X_3}, \frac{X_3}{X_1 + X_2 + X_3}\right).$$

(a) Show that $Y \sim \text{Dirichlet}(a_1, a_2, a_3)$, a Dirichlet distribution.

First, we consider the joint density of the three independent Gamma-distributed RVs:

$$p(x_1, x_2, x_3) = \prod_{i=1}^{3} p(x_i) = \prod_{i=1}^{3} \frac{x_i^{a_i - 1} e^{-x_i/b}}{\Gamma(a_i) b^{a_i}} = \frac{e^{\sum_{i=1}^{3} x_i/b} \prod_{i=1}^{3} x_i^{a_i - 1}}{b^{\sum_{i=1}^{3} a_i} \prod_{i=1}^{3} \Gamma(a_i)}.$$

Since $Y_i = X_i / \sum_{i=1}^3 X_i$, we find

$$X_1 = Y_1 Z$$

 $X_2 = Y_2 Z$
 $X_3 = Y_3 Z = (1 - Y_1 - Y_2) Z$,

where $Z = \sum_{i=1}^{3} X_i$. To obtain the joint distribution of (Y_1, Y_2, Z) , we find the Jacobian for this change of variables. The matrix is

$$J = \begin{bmatrix} \frac{dx_1}{dy_1} & \frac{dx_1}{dy_2} & \frac{dx_1}{dz} \\ \frac{dx_2}{dy_1} & \frac{dx_2}{dy_2} & \frac{dx_2}{dz} \\ \frac{dx_3}{dy_1} & \frac{dx_3}{dy_2} & \frac{dx_3}{dz} \end{bmatrix} = \begin{pmatrix} Z & 0 & Y_1 \\ 0 & Z & Y_2 \\ -Z & -Z & (1 - Y_1 - Y_2) \end{pmatrix}$$

So, the Jacobian, |J| is Z^2 .

$$p(Y_1, Y_2, Z) = \frac{(y_1 z)^{a_1 - 1} e^{-y_1 z/b} (y_2 z)^{a_2 - 1} e^{-y_2 z/b} \{ (1 - y_1 - y_2) z \}^{a_3 - 1} e^{-(1 - y_1 - y_2) z/b}}{b^{\sum_{i=1}^3 a_i} \prod_{i=1}^3 \Gamma(a_i)} z^2$$

where $0 < y_1, y_2 < 1, y_1 + y_2 < 1$ and 0 < z.

By letting $y_3 = 1 - y_1 - y_2$,

$$p(Y_1, Y_2, Z) = \frac{y_1^{a_1 - 1} y_2^{a_2 - 1} y_3^{a_3 - 1} z^{\sum_{i=1}^{3} a_i - 1} e^{z} \sum_{i=1}^{3} y_i / b}{b^{\sum_{i=1}^{3} a_i} \prod_{i=1}^{3} \Gamma(a_i)}.$$

We now integrate out z to obtain $p(y_1, y_2, y_3)$.

kernel for Gamma(
$$\sum_{i=1}^{3} a_i, b$$
)

$$p(y_1, y_2, y_3) = \int_{\mathbb{R}} \frac{y_1^{a_1 - 1} y_2^{a_2 - 1} y_3^{a_3 - 1}}{b^{\sum_{i=1}^3 a_i} \prod_{i=1}^3 \Gamma(a_i)} dz$$

$$= \frac{\Gamma(\sum_{i=1}^3 a_i)}{\prod_{i=1}^3 \Gamma(a_i)} y_1^{a_1 - 1} y_2^{a_2 - 1} y_3^{a_3 - 1}.$$

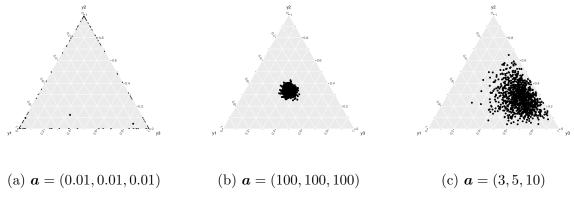


Figure 1: $\mathbf{y} = (y_1, y_2, y_3)$ simulated from Dirichlet distribution $\mathrm{Dir}(\mathbf{a})$.

```
Thus, Y = (Y_1, Y_2, Y_3) \sim \text{Dir}(a_1, a_2, a_3).
```

(b) How can this result be used to generate random variables according to a Dirichlet distribution? Write a simple function in R or Matlab (your choice) that takes as inputs n, the number of trivariate vectors to be generated, and $\mathbf{a} = (a_1, a_2, a_3)$ and generates a matrix of size $n \times 3$ whose rows correspond to independent samples from a Dirichlet distribution with parameter (a_1, a_2, a_3) .

Use each of $\mathbf{a} = (0.01, 0.01, 0.01)$, (100, 100, 100), and (3, 5, 10) and comment how the density of \mathbf{Y} changes over \mathbf{a} .

```
#a is a vector of length p; a=(a_1, a_2, ..., a_p)
#n is the sample size

> dirichlet <- function(a, n){
        p <- length(a)
        y <- array(NA, dim=c(n, p)) #Each row of y is iid sample from Dir(a)

        for (i in 1:n) {
            tmp <- rgamma(p, a, 1)
            y[i, ] <- tmp / sum(tmp)
        }

        return(y)
    }
</pre>
```

Fig 1 illustrates trinary plots of 1000 simulated y with different a. For a < 1, y are at the vertices and edges, meaning either one or two of y_1, y_2 or y_3 are close to zero. For a = 100, all y_i are close to 1/3 (i.e., small variance) as shown in Fig 1(b). For a = (3, 5, 10), y are around the mean (3/18, 5/18, 10/18) and its variance is larger than that with a = 100.

- 4. Y follows an inverse Gamma distribution with shape parameter a and scale parameter b $(Y \sim \mathrm{IG}(a,b))$ if Y=1/X with $X \sim \mathrm{Gamma}(a,b)$ (assume the Gamma distribution is parameterized so that $\mathrm{E}(X)=ab$.
 - (a) Find the density of Y.

Since $X \sim \text{Gamma}(a, b)$,

$$f(x) = \frac{b^a}{\Gamma(a)} x^{a-1} \exp(-bx), \text{ for } x > 0.$$

Let y = 1/x. Then x = 1/y and $\frac{dx}{dy} = -1/y^2$. Therefore,

$$p(y) = \frac{b^a}{\Gamma(a)} \left(\frac{1}{y}\right)^{a-1} \exp(-b/y) \left| -\frac{1}{y^2} \right| = \frac{b^a}{\Gamma(a)} y^{-(a+1)} \exp(-b/y), \ y > 0,$$

which is an inverse Gamma with shape a and rate $b, y \sim IG(a, b)$.

(b) Compute $E(Y^k)$. Do you need to impose any constrain on the problem for this expectation to exists?

We have

$$E(Y^k) = \int_{\mathbb{R}^+} y^k \frac{b^a}{\Gamma(a)} y^{-(a+1)} \exp(-b/y) dy = \frac{b^a}{\Gamma(a)} \frac{\Gamma(a-k)}{b^{a-k}} = \frac{b^k \Gamma(a-k)}{\Gamma(a)}.$$

That is, $E(Y^k) = b^k \Gamma(a-k)/\Gamma(a)$ for a-k > 0.

(c) Compare $E(Y^k)$ to $1/E(X^k)$ (hint: look at the ratio of the two quantities).

We first find

$$\mathrm{E}(X^k) = \frac{b^a \Gamma(a+k)}{\Gamma(a) b^{a+k}} = \frac{\Gamma(a+k)}{\Gamma(a) b^k}.$$

Then,

$$\zeta \equiv \frac{1/\mathrm{E}(X^k)}{\mathrm{E}(Y^k)} = \frac{\Gamma(a)\Gamma(a)}{\Gamma(a+k)\Gamma(a-k)}$$

which implies that $\zeta = 1 \Leftrightarrow k = 0$. That is, none of the moments of X is invariant to the reciprocal transformation.

5. Y follows a log normal distribution with parameters μ and σ^2 (denotes as $Y \sim \text{Log-N}(\mu, \sigma^2)$ if $Y = \exp(X)$ where $X \sim \text{N}(\mu, \sigma^2)$).

4

(a) Find the density of Y.

We have $Y = \exp(X) \in \mathbb{R}^+$. We use a change of a variable and find

$$g(y) = f(\log(y)) \left| \frac{1}{y} \right| = \frac{1}{\sqrt{2\pi\sigma^2}y} \exp\left\{ -\frac{(\log(y) - \mu)^2}{2\sigma^2} \right\}, \ y \in \mathbb{R}^+.$$

(b) Compute the mean and the variance of Y.

Observe the k-th moment of Y, $E^{Y}(Y^{k}) = E^{X}(e^{kX})$ and find

$$E^{X}(e^{kX}) = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi\sigma^{2}}} \exp\left\{-\frac{(x-\mu)^{2}}{2\sigma^{2}} + kx\right\} dx$$
$$= \exp\left\{-\frac{\mu^{2}}{2\sigma^{2}} + \frac{(\mu+k\sigma^{2})^{2}}{2\sigma^{2}}\right\}$$
$$= \exp\left(k\mu + k^{2}\sigma^{2}/2\right).$$

Thus, $E(Y) = \exp(\mu + \sigma^2/2)$ and $Var(Y) = E(Y^2) - E(Y)^2 = \exp(2\mu + \sigma^2) \{\exp(\sigma^2) - 1\}$.

6. Let $\boldsymbol{X}=(X_1,X_2,\ldots,X_p)$ with $X\sim \mathrm{N}_p(\boldsymbol{\mu},\Sigma)$ and set $\boldsymbol{Z}_1=(X_1,\ldots,X_q)$ and $\boldsymbol{Z}_2=(X_{q+1},\ldots,X_p)$ with 1< q< p. Show that

$$Z_1 \mid Z_2 \sim N_q \left(\mu_1 + \Sigma_{12} \Sigma_{22}^{-1} (Z_2 - \mu_2), \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \right),$$

where μ_k and $\Sigma_{k\ell}$ denote the blocks of μ and Σ where the rows correspond to the variables in \mathbb{Z}_k and the columns to the variables in \mathbb{Z}_ℓ .

Using some results of the inverse of a partitioned matrix, we find for symmetric and $\Sigma > 0$,

$$\Sigma^{-1} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \Sigma_{11.2}^{-1} & -\Sigma_{11.2}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \\ -\Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11.2}^{-1} & \Sigma_{22.1}^{-1} \end{bmatrix} = \begin{bmatrix} \Sigma_{11.2}^{-1} & -\Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22.1}^{-1} \\ -\Sigma_{22.1}^{-1} \Sigma_{21} \Sigma_{21}^{-1} & \Sigma_{22.1}^{-1} \end{bmatrix},$$

where $\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$ and $\Sigma_{22.2} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$. We have

$$\begin{split} g(\boldsymbol{z}_1 \mid \boldsymbol{z}_2) &\propto f(\boldsymbol{z}_1, \boldsymbol{z}_2) \\ &\propto \exp\left\{-\frac{1}{2}\begin{bmatrix}\boldsymbol{z}_1 - \boldsymbol{\mu}_1 \\ \boldsymbol{z}_2 - \boldsymbol{\mu}_2\end{bmatrix}' \boldsymbol{\Sigma}^{-1}\begin{bmatrix}\boldsymbol{z}_1 - \boldsymbol{\mu}_1 \\ \boldsymbol{z}_2 - \boldsymbol{\mu}_2\end{bmatrix}\right\} \\ &\propto \exp\left[-\frac{1}{2}\Big\{(\boldsymbol{z}_1 - \boldsymbol{\mu}_1)'\boldsymbol{\Sigma}_{11.2}^{-1}(\boldsymbol{z}_1 - \boldsymbol{\mu}_1) - (\boldsymbol{z}_1 - \boldsymbol{\mu}_1)'\boldsymbol{\Sigma}_{11.2}^{-1}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{z}_2 - \boldsymbol{\mu}_2) \right. \\ &\left. - (\boldsymbol{z}_2 - \boldsymbol{\mu}_2)'\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11.2}^{-1}(\boldsymbol{z}_1 - \boldsymbol{\mu}_1)\Big\}\right] \\ &\propto \exp\left[-\frac{1}{2}\Big\{\boldsymbol{z}_1'\boldsymbol{\Sigma}_{11.2}^{-1}\boldsymbol{z}_1 - \boldsymbol{z}_1'\left(\boldsymbol{\Sigma}_{11.2}^{-1}\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{11.2}^{-1}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{z}_2 - \boldsymbol{\mu}_2)\right) \right. \\ &\left. - (\boldsymbol{\mu}_1'\boldsymbol{\Sigma}_{11.2}^{-1} + (\boldsymbol{z}_2 - \boldsymbol{\mu}_2)'\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11.2}^{-1})\boldsymbol{z}_1\Big\}\right]. \end{split}$$

We recognize the kernel for $N_q(\Sigma_{11.2} \left(\Sigma_{11.2}^{-1} \boldsymbol{\mu}_1 + \Sigma_{11.2}^{-1} \Sigma_{12} \Sigma_{22}^{-1} (\boldsymbol{z}_2 - \boldsymbol{\mu}_2)\right), \Sigma_{11.2}).$ That is, $\boldsymbol{Z}_1 \mid \boldsymbol{Z}_2 \sim N_q \left(\boldsymbol{\mu}_1 + \Sigma_{12} \Sigma_{22}^{-1} (\boldsymbol{Z}_2 - \boldsymbol{\mu}_2), \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}\right).$

7. Show that if $X \sim \text{Exp}(\beta)$, then

(a) $Y = X^{1/\gamma}$ has a Weibull distribution with parameters γ and β with $\gamma > 0$ a constant.

We have $Y = X^{1/\gamma} \in \mathbb{R}^+$. We use a change of a variable and find

$$g(y) = f(y^{\gamma})|\gamma y^{\gamma-1}| = \beta \gamma y^{\gamma-1} \exp(-\beta y^{\gamma}), \quad y \in \mathbb{R}^+.$$

We also observe $\mathrm{E}^Y(Y^k) = \mathrm{E}^X(X^{k/\gamma}) = \frac{\Gamma(1+k/\gamma)}{\beta^{k/\gamma}}$ (use the gamma kernel). We find $\mathrm{E}^Y(Y) = \frac{\Gamma(1+1/\gamma)}{\beta^{1/\gamma}}$ and $\mathrm{Var}(Y) = \mathrm{E}^X(X^{2/\gamma}) - \mathrm{E}^X(X^{1/\gamma})^2 = \left\{\Gamma(1+2/\gamma) - (\Gamma(1+1/\gamma))^2\right\}/\beta^{2/\gamma}$.

(b) $Y = (2X/\beta)^{1/2}$ has the Rayleigh distribution.

We have $Y = (2X/\beta)^{1/2} \in \mathbb{R}^+$. We use a change of a variable and find

$$g(y) = f(\beta y^2/2)|\beta y| = \beta^2 y \exp(-\beta^2 y^2/2), \ y \in \mathbb{R}^+.$$

We also observe $E^Y(Y^k) = E^X((2X/\beta)^{k/2}) = \frac{2^{k/2}\Gamma(1+k/2)}{\beta^k}$ (use the gamma kernel). We find $E^Y(Y) = \sqrt{2}\Gamma(1.5)/\beta$ and $Var(Y) = E^X(2X/\beta) - E^X((2X/\beta)^{1/2})^2 = 2(1 - \Gamma(1.5)^2)/\beta^2$.

For both parts, derive the form of the pdf, verify that is a pdf, and calculate the mean and the variance.

8. Let $Y \mid X \sim \text{Poisson}(X)$ and let $X \sim \text{Exp}(\lambda)$. What is the marginal distribution of Y?

We find

$$m(y) = \int_{\mathbb{R}^+} \frac{e^{-x} x^y}{y!} \lambda e^{-\lambda x} dx = \frac{\Gamma(y+1)\lambda}{y!(1+\lambda)^{y+1}} = \frac{\lambda}{(1+\lambda)^{y+1}}, y = 0, 1, 2, \dots$$

9. (Robert 1.37) Let $x \sim N(\theta, \sigma^2)$, $y \sim N(\rho x, \sigma^2)$ with ρ known. Assume a prior of the form $\pi(\theta, \sigma^2) = 1/\sigma^2$. Find the predictive density of y given x.

We check the marginal distribution $\int_{\mathbb{R}} m(x)dx < 0$.

$$\int_{\mathbb{R}} m(x)dx = \int_{\mathbb{R}^+} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-\theta)^2}{2\sigma^2}\right) (\sigma^2)^{-1/2-1} d\theta d\sigma^2$$
$$= \int_{\mathbb{R}^+} (\sigma^2)^{-1} d\sigma^2 = \infty.$$

Thus, $\theta, \sigma^2 \mid x$ cannot be defined and $y \mid x$ cannot.

10. (Robert) If $y \sim \text{Binomial}(n, \theta)$ and $x \sim \text{Binomial}(m, \theta)$, and $\theta \sim \text{Beta}(\alpha, \beta)$. Find the predictive distribution of y given x.

Assume conditional independence of X and Y given θ . We know that $\theta \mid x \sim \text{Be}(\alpha + x, \beta + m - x)$ and then have for $y = 0, \dots, n$,

$$f(y \mid x) = \int_0^1 f(y \mid \theta) \pi(\theta \mid x) d\theta$$

$$= \int_0^1 \binom{n}{y} \theta^y (1 - \theta)^{n-y} \frac{\theta^{\alpha+x-1} (1 - \theta)^{\beta+m-x-1}}{B(\alpha + x, \beta + m - x)} d\theta$$

$$= \frac{\binom{n}{y} B(\alpha + x + y, \beta + m - x + n - y)}{B(\alpha + x, \beta + m - x)}, \quad y = 0, 1, 2, \dots$$

that is, $y \mid x \sim \text{Beta-Binomial}(n, \alpha + x, \beta + m - x)$.

- 11. (Robert) Give the posterior and the marginal distributions in the following cases:
 - (a) $x \mid \sigma^2 \sim N(0, \sigma^2)$ and $1/\sigma^2 \sim Gamma(1, 2)$.

We know $1/\sigma^2 \sim \text{Gamma}(1,2) \implies \sigma^2 \sim \text{IG}(1,2)$.

$$m(x) = \int_0^\infty \frac{1}{\sqrt{2\pi}} (\sigma^2)^{-1/2} \exp\left(-\frac{x^2}{2\sigma^2}\right) \frac{2}{\Gamma(1)} (\sigma^2)^{-2} \exp\left(-\frac{2}{\sigma^2}\right) d\sigma^2$$
$$= \frac{2\Gamma(3/2)}{\sqrt{2\pi} (2 + x^2/2)^{3/2}}, \quad x \in \mathbb{R}.$$

We have

$$\pi(\sigma^2\mid x) \propto (\sigma^2)^{-1/2} \exp\left(-\frac{x^2}{2\sigma^2}\right) (\sigma^2)^{-2} \exp\left(-\frac{2}{\sigma^2}\right), 0 < \sigma^2$$

that is, $\sigma^2 \mid x \sim \text{IG}(3/2, 2 + x^2/2) \ (\Rightarrow 1/\sigma^2 \mid x \sim \text{Gamma}(3/2, 2 + x^2/2)).$

(b) $x \mid p \sim \text{Negative-Binomial}(10, p)$ and $p \sim \text{Beta}(1/2, 1/2)$.

Following the parameterization of CR p522, we have

$$m(x) = \int_0^\infty {n+x+1 \choose x} p^n (1-p)^x \frac{1}{B(1/2,1/2)} p^{-1/2} (1-p)^{-1/2} dp$$
$$= \frac{{n+x+1 \choose x} B(n+1/2,x+1/2)}{B(1/2,1/2)}, x = 0, 1, \dots$$

We have

$$\pi(p \mid x) \propto p^n (1-p)^x p^{-1/2} (1-p)^{-1/2}, 0$$

that is, $p \mid x \sim \text{Be}(n + 1/2, x + 1/2)$.