Petter Mostad
Applied Mathematics and Statistics
Chalmers and GU

MVE550 Stochastic Processes and Bayesian Inference

Re-exam August 21, 2023, 8:30 - 12:30
Examiner: Petter Mostad will be available by phone 031-772-3579
Allowed aids: Chalmers-approved calculator
Total number of points: 30. At least 12 points are needed to pass.
See appendix for some information about some probability distributions
All answers need to be explicitly computed or explicitly argued for.

1. (1 point) Define what it means for a stochastic process to have the Markov property: Define it both in words, and in a formula for a process $\{X_t\}_{t\in I}$ where t indicates time.

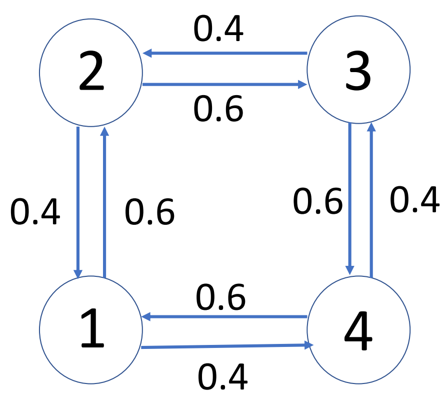

Figure 1: The transition graph for question 2.

2. (4 points)

   (a) What are the conditions for a discrete-time discrete state space Markov chain to be ergodic?

   (b) Please describe (precisely) the most important limiting theorem that holds for such ergodic Markov chains.

   (c) Is the Markov chain with the transition graph in Figure 1 ergodic? Please explain.

(d) Let P be the transition matrix of the Markov chain of Figure 1. What[1] is the value of

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=0}^{n-1} P^m_{44}$$

3. (8 points) A drop-in bike workshop has two employees, Hans and Otto. The time it takes for Hans to fix a bike is 30 minutes on average, while Otto uses 1 hour on average. Customers arrive independently of each other at a rate of 2 per hour. If neither Otto nor Hans is working when a customer arrive, they flip a coin to decide who does the repair. If they are both working, a new customer waits in line. If there are two persons waiting already, new customers walk away. You may assume repair times are exponentially distributed.

   (a) Modelling the workshop with a continuous-time Markov chain, draw a transition rate diagram for the chain. Also, write down the generator matrix.

   (b) What is the long-term proportion of time that Otto is working on fixing a bike? Instead of computing the result as a number, you may express the result as an equation containing vectors and matrices of numbers.

   (c) If they start working at 9 in the morning, what is the expected number of hours until a customer walks away because there are already 2 persons waiting? Instead of computing the result as a number, you may express the result as an equation containing vectors and matrices of numbers.

   (d) Is this a birth-and-death process? Please argue yes or no.

4. (5 points) Pedro makes independent observations of random variables X that are Gamma(α,β) distributed, where $\alpha > 0$ and $\beta > 0$ are two unknown parameters. He has observed $x_1 = 2.4$, $x_2 = 3.1$, and $x_3 = 2.9$. He would now like to make a probabilistic prediction for x_4, a fourth observation, using the information learned in the three first observations.

   (a) What is the likelihood function for (α,β) given this data? Write down and simplify.

   (b) Pedro first assumes an prior *(misprint in original exam: "imprior")* density

$$\pi(\alpha,\beta) \propto \begin{cases} 1/\beta & \text{if } 4 \le \alpha \le 7 \\ 0 & \text{otherwise} \end{cases}$$

   Write down a function proportional to the posterior density for (α,β).

   (c) Describe one possible numerical computation with which Pedro can compute the posterior probability that $x_4 > 3$.

   (d) Now, instead of the prior from (a), Pedro uses a prior where $\alpha = 5$ is fixed, and $\pi(\beta) \propto 1/\beta$. Write down in an analytical form the posterior density for β in this case.

[1]Recall that P^m_{44} should be interpreted as $(P^m)_{44}$.

5. (5 points) Let X be a random variable with the non-negative integers as state space.

   (a) Define the probability generating function $G_X(s)$.

   (b) Compute $G'_X(1)$ and $G''_X(X)$ in terms of properties of X.

   (c) State and prove a way to express $\mathrm{Var}[X]$ in terms of $G'_X(1)$ and $G''_X(1)$.

6. (3 points) Assume $\{B_t\}_{t\geq 0}$ is Brownian motion and that $k > 1$ is an integer. Compute the distribution of

$$B_1 + B_2 + \cdots + B_k.$$

7. (4 points) Assume given a generator matrix Q for a continuous-time Markov chain. Poisson subordination means that one introduces a matrix $R = \frac{1}{\lambda}Q + I$ for some λ.

   (a) What is the condition on λ for R to become a stochastic matrix? Prove your answer.

   (b) Show how the exponential matrix e^{tQ} can be expressed as an infinite sum expressed in matrix powers R^k.

   (c) Explain how your formula from (b) can be used to make numerical approximate computations for e^{tQ}.

Appendix: Some probability distributions

The Beta distribution

If $x \in [0, 1]$ has a Beta distribution with parameters with $\alpha > 0$ and $\beta > 0$ then the density is

$$\pi(x \mid \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1 - x)^{\beta-1}.$$

We write $x \mid \alpha, \beta \sim \text{Beta}(\alpha, \beta)$ and $\pi(x \mid \alpha, \beta) = \text{Beta}(x; \alpha, \beta)$.

The Beta-Binomial distribution

If $x \in \{0, 1, 2, \ldots, n\}$ has a Beta-Binomial distribution, with n a positive integer and parameters $\alpha > 0$ and $\beta > 0$, then the probability mass function is

$$\pi(x \mid n, \alpha, \beta) = \binom{n}{x} \frac{\Gamma(x + \alpha)\Gamma(n - x + \beta)\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(n + \alpha + \beta)}.$$

We write $x \mid n, \alpha, \beta \sim \text{Beta-Binomial}(n, \alpha, \beta)$ and $\pi(x \mid n, \alpha, \beta) = \text{Beta-Binomial}(x; n, \alpha, \beta)$.

The Binomial distribution

If $x \in \{0, 1, 2, \ldots, n\}$ has a Binomial distribution, with n a positive integer and $0 \leq p \leq 1$, then the probability mass function is

$$\pi(x \mid n, p) = \binom{n}{x} p^x (1 - p)^{n-x}.$$

We write $x \mid n, p \sim \text{Binomial}(n, p)$ and $\pi(x \mid n, p) = \text{Binomial}(x; n, p)$.

The Dirichlet distribution

If $x = (x_1, x_2, \ldots, x_n)$ has a Dirichlet distribution, with $x_i \geq 0$ and $\sum_{i=1}^n x_i = 1$ and with parameters $\alpha = (\alpha_1, \ldots, \alpha_n)$ with $\alpha_1 > 0, \ldots, \alpha_n > 0$, then the density function is

$$\pi(x \mid \alpha) = \frac{\Gamma(\alpha_1 + \alpha_2 + \cdots + \alpha_n)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\cdots\Gamma(\alpha_n)} p_1^{\alpha_1-1} p_2^{\alpha_2-1} \cdots p_n^{\alpha_n-1}.$$

We write $x \mid \alpha \sim \text{Dirichlet}(\alpha)$ and $\pi(x \mid \alpha) = \text{Dirichlet}(x; \alpha)$.

The Exponential distribution

If $x \geq 0$ has an Exponential distribution with parameter $\lambda > 0$, then the density is

$$\pi(x \mid \lambda) = \lambda \exp(-\lambda x)$$

We write $x \mid \lambda \sim \text{Exponential}(\lambda)$ and $\pi(x \mid \lambda) = \text{Exponential}(x; \lambda)$. The expectation is $1/\lambda$ and the variance is $1/\lambda^2$.

The Gamma distribution

If $x > 0$ has a Gamma distribution with parameters $\alpha > 0$ and $\beta > 0$ then the density is

$$\pi(x \mid \alpha\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x).$$

We write $x \mid \alpha, \beta \sim \text{Gamma}(\alpha, \beta)$ and $\pi(x \mid \alpha, \beta) = \text{Gamma}(x; \alpha, \beta)$.

The Geometric distribution

If $x \in \{1, 2, 3, \ldots\}$ has a Geometric distribution with parameter $p \in (0, 1)$, the probability mass function is

$$\pi(x \mid p) = p(1-p)^{x-1}$$

We write $x \mid p \sim \text{Geometric}(p)$ and $\pi(x \mid p) = \text{Geometric}(x; p)$. The expectation is $1/p$ and the variance $(1-p)/p^2$.

The Negative Binomial distribution

A stochastic variable x taking on as possible values any nonnegative integer has a Negative Binomial distribution if its probability mass function is given by

$$\pi(x \mid r, p) = \binom{x+r-1}{x} \cdot (1-p)^x p^r = \frac{\Gamma(x+r)}{\Gamma(x+1)\Gamma(r)} (1-p)^x p^r$$

where $r > 0$ and $p \in (0, 1)$ are parameters.

The Normal distribution

If the real x has a Normal distribution with parameters μ and σ^2, its density is given by

$$\pi(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right).$$

We write $x \mid \mu, \sigma^2 \sim \text{Normal}(\mu, \sigma^2)$ and $\pi(x \mid \mu, \sigma^2) = \text{Normal}(x; \mu, \sigma^2)$.

The Poisson distribution

If $x \in \{0, 1, 2, \ldots\}$ has Poisson distribution with parameter $\lambda > 0$ then the probability mass function is

$$e^{-\lambda}\frac{\lambda^x}{x!}.$$

We write $x \mid \lambda \sim \text{Poisson}(\lambda)$ and $\pi(x \mid \lambda) = \text{Poisson}(x; \lambda)$. The Poisson distribution has expectation λ and variance λ.

Petter Mostad
Applied Mathematics and Statistics
Chalmers

Suggested solutions for
MVE550 Stochastic Processes and Bayesian Inference
Re-exam August 21 2023

1. In words: The distribution of the process in the future is independent of its past given its value in the present. In a formula: For all $r \in I$,

$$\Pr[X_s, s > r \mid X_t, t \leq r] = \Pr[X_s, s > r \mid X_r].$$

2. (a) A discrete-time discrete state space Markov chain is ergodic if it is irreducible, positive recurrent, and aperiodic.

   (b) For such chains the "Fundamental Limit Theorem" holds, stating that there **exists** a **unique positive stationary** distribution with is also the **limiting** distribution for the chain.

   (c) No, it is not ergodic. In fact it is periodic with period 2, as the chain will always jump between even and odd states.

   (d) Even if the chain is not ergodic, it is finite and irreducible, and then the theorem for such chains states that there is a unique positive stationary distribution for the chain whose j'th term is

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=0}^{n-1} P_{ij}^m$$

   for any i. Thus the value we want is the 4'th term of the unique stationary distribution for the chain. Looking at Figure 1 of the questions document, we see that the chain is symmetric in the four states, so $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ must be a stationary distribution, and thus the unique stationary distribution. Thus the answer is $\frac{1}{4}$.

3. (a) Naming the 6 necessary states and listing the states in the following order

   - I: Both persons are idle.
   - H: Hans works while Otto is idle.
   - O: Otto works while Hans is idle.
   - B0W: Both work, zero persons waiting.
   - B1W: Both work, one person waiting.
   - B2W: Both work, two persons waiting.

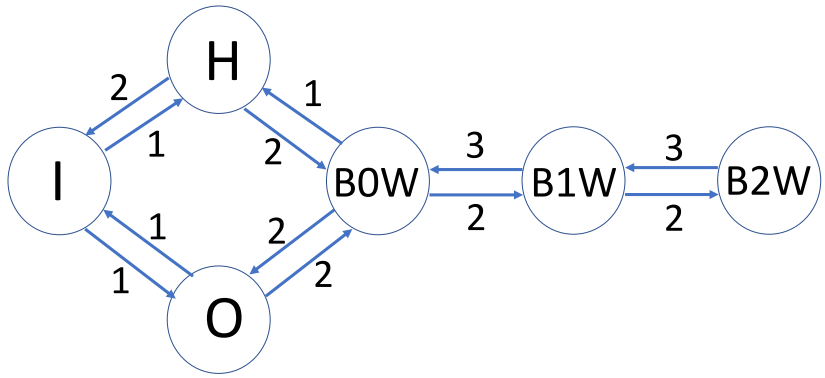

Figure 1: The transition rate graph for question 3.

we get the diagram of Figure 1 and the generator matrix

$$Q = \begin{bmatrix} -2 & 1 & 1 & 0 & 0 & 0 \\ 2 & -4 & 0 & 2 & 0 & 0 \\ 1 & 0 & -3 & 2 & 0 & 0 \\ 0 & 1 & 2 & -5 & 2 & 0 \\ 0 & 0 & 0 & 3 & -5 & 2 \\ 0 & 0 & 0 & 0 & 3 & -3 \end{bmatrix}.$$

(b) This is an irreducible chain, so it has a limiting distribution v which can be found by solving the equation $vQ = 0$ for a vector of positive values summing to 1. We can find this vector by solving

$$v \begin{bmatrix} -2 & 1 & 1 & 0 & 0 & 1 \\ 2 & -4 & 0 & 2 & 0 & 1 \\ 1 & 0 & -3 & 2 & 0 & 1 \\ 0 & 1 & 2 & -5 & 2 & 1 \\ 0 & 0 & 0 & 3 & -5 & 1 \\ 0 & 0 & 0 & 0 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Otto works when the chain is in state O, B0W, B1W, or B2W. This means that the proportion of time Otto works can be computed as the sum of the four last components of v. In other words, the proportion can be computed as

$$\begin{bmatrix}0 & 0 & 0 & 0 & 0 & 1\end{bmatrix}\begin{bmatrix}-2 & 1 & 1 & 0 & 0 & 1\\ 2 & -4 & 0 & 2 & 0 & 1\\ 1 & 0 & -3 & 2 & 0 & 1\\ 0 & 1 & 2 & -5 & 2 & 1\\ 0 & 0 & 0 & 3 & -5 & 1\\ 0 & 0 & 0 & 0 & 3 & 1\end{bmatrix}^{-1}\begin{bmatrix}0\\0\\1\\1\\1\\1\end{bmatrix}.$$

(c) We can solve this by imagining that there is a seventh state B3W which is absorbing. The generator matrix would then become

$$Q = \begin{bmatrix}-2 & 1 & 1 & 0 & 0 & 0 & 0\\ 2 & -4 & 0 & 2 & 0 & 0 & 0\\ 1 & 0 & -3 & 2 & 0 & 0 & 0\\ 0 & 1 & 2 & -5 & 2 & 0 & 0\\ 0 & 0 & 0 & 3 & -5 & 2 & 0\\ 0 & 0 & 0 & 0 & 3 & -5 & 2\\ 0 & 0 & 0 & 0 & 0 & 0 & 0\end{bmatrix}$$

and we get the corresponding fundamental matrix

$$F = -\begin{bmatrix}-2 & 1 & 1 & 0 & 0 & 0\\ 2 & -4 & 0 & 2 & 0 & 0\\ 1 & 0 & -3 & 2 & 0 & 0\\ 0 & 1 & 2 & -5 & 2 & 0\\ 0 & 0 & 0 & 3 & -5 & 2\\ 0 & 0 & 0 & 0 & 3 & -5\end{bmatrix}^{-1}.$$

Finally, the expected number of hours until absorbtion when the chain starts in state I can now be computed as the sum of the first row of F:

$$-\begin{bmatrix}1 & 0 & 0 & 0 & 0 & 0\end{bmatrix}\begin{bmatrix}-2 & 1 & 1 & 0 & 0 & 0\\ 2 & -4 & 0 & 2 & 0 & 0\\ 1 & 0 & -3 & 2 & 0 & 0\\ 0 & 1 & 2 & -5 & 2 & 0\\ 0 & 0 & 0 & 3 & -5 & 2\\ 0 & 0 & 0 & 0 & 3 & -5\end{bmatrix}^{-1}\begin{bmatrix}1\\1\\1\\1\\1\\1\end{bmatrix}.$$

(d) This is NOT a birth-and-death process. The transition graph of a birth-and-death process is linear, while the transition graph of this process is not.

4. (a) The likelihood function becomes

$$\begin{aligned}
\pi(x_1, x_2, x_3 \mid \alpha, \beta) &= \prod_{i=1}^{3} \text{Gamma}(x_i; \alpha, \beta) \\
&= \prod_{i=1}^{3} \frac{\beta^\alpha}{\Gamma(\alpha)} x_i^{\alpha-1} \exp(-\beta x_i) \\
&= \frac{\beta^{3\alpha}}{\Gamma(\alpha)^3} (x_1 x_2 x_3)^{\alpha-1} \exp(-\beta(x_1 + x_2 + x_3)) \\
&= \frac{\beta^{3\alpha}}{\Gamma(\alpha)^3} 25.576^{\alpha-1} \exp(-8.4\beta)
\end{aligned}$$

(b) The posterior can be computed, up to proportionality, as the product

$$\begin{aligned}
\pi(\alpha, \beta \mid x_1, x_2, x_3) &\propto \pi(x_1, x_2, x_3 \mid \alpha, \beta)\pi(\alpha, \beta) \\
&\propto \frac{\beta^{3\alpha-1}}{\Gamma(\alpha)^3} 25.576^{\alpha-1} \exp(-8.4\beta)
\end{aligned}$$

whenever $4 \leq \alpha \leq 7$, and zero otherwise.

(c) We can use numerical integration or discretization. More specifically, we can write

$$\begin{aligned}
\pi(x_4 > 3 \mid x_1, x_2, x_3) &= \int_{\beta=0}^{\infty} \int_{\alpha=4}^{7} \pi(x_4 > 3 \mid \alpha, \beta)\pi(\alpha, \beta \mid x_1, x_2, x_3)\, d\alpha\, d\beta \\
&= \frac{\int_{\beta=0}^{\infty} \int_{\alpha=4}^{7} \pi(x_4 > 3 \mid \alpha, \beta)\pi(x_1, x_2, x_3 \mid \alpha, \beta)\pi(\alpha, \beta)\, d\alpha\, d\beta}{\int_{\beta=0}^{\infty} \int_{\alpha=4}^{7} \pi(x_1, x_2, x_3 \mid \alpha, \beta)\pi(\alpha, \beta)\, d\alpha\, d\beta}
\end{aligned}$$

These integrals can be computed plugging in the expression from (b) and noting that $\pi(x_4 > 3 \mid \alpha, \beta)$ is simply the value of the cumulative Gamma distribution, which may be computed in R with the function `pgamma`.

(d) Substituting $\alpha = 5$ into the expression for the posterior found in (b) we get

$$\begin{aligned}
\pi(\beta \mid x_1, x_2, x_3) &\propto \frac{\beta^{3\cdot 5-1}}{\Gamma(5)^3} 25.576^{5-1} \exp(-8.4\beta) \\
&\propto \beta^{15-1} \exp(-8.4\beta) \\
&\propto \frac{15^{8.4}}{\Gamma(8.4)} \beta^{15-1} \exp(-8.4\beta) \\
&= \text{Gamma}(\beta; 15, 8.4).
\end{aligned}$$

As the posterior is proportional to the Gamma density above, it must be equal to the Gamma density above.

5. (a) $G_X(s) = \mathrm{E}\left[s^X\right]$.

(b) We may compute

$$\begin{aligned}
G'_X(s) &= \frac{d}{ds}\,\mathrm{E}\left[s^X\right] \\
&= \frac{d}{ds}\sum_{k=0}^{\infty} s^k \Pr[X=k] \\
&= \sum_{k=0}^{\infty} \frac{d}{ds} s^k \Pr[X=k] \\
&= \sum_{k=0}^{\infty} k s^{k-1} \Pr[X=k] \\
&= \mathrm{E}\left[X s^{X-1}\right]
\end{aligned}$$

so $G'_X(1) = \mathrm{E}\left[X 1^{X-1}\right] = \mathrm{E}[X]$. We also have

$$\begin{aligned}
G''_X(s) &= \frac{d}{ds}\sum_{k=0}^{\infty} k s^{k-1} \Pr[X=k] \\
&= \sum_{k=0}^{\infty} \frac{d}{ds} k s^{k-1} \Pr[X=k] \\
&= \sum_{k=0}^{\infty} k(k-1) s^{k-2} \Pr[X=k] \\
&= \mathrm{E}\left[X(X-1)s^{X-2}\right]
\end{aligned}$$

so $G''_X(1) = \mathrm{E}\left[X(X-1)1^{X-2}\right] = \mathrm{E}[X(X-1)] = \mathrm{E}\left[X^2\right] - \mathrm{E}[X]$.

(c) We may now write

$$\begin{aligned}
\mathrm{Var}[X] &= \mathrm{E}\left[X^2\right] - \mathrm{E}[X]^2 \\
&= \mathrm{E}\left[X^2\right] - \mathrm{E}[X] + \mathrm{E}[X] - \mathrm{E}[X]^2 \\
&= G''_X(1) + G'_X(1) - G'_X(1)^2.
\end{aligned}$$

6. Building on the examples from the course we can fairly easily show that

$$B_1 + \cdots + B_k = kB_1 + (k-1)(B_2 - B_1) + (k-2)(B_3 - B_2) + \cdots + B_k - B_{k-1}.$$

As the variables $B_i - B_{i-1}$ are all independent and Normal(0, 1) distributed, we get that the sum above also has a Normal distribution. Its expectation is

$$\mathrm{E}[B_1 + \cdots + B_k] = \mathrm{E}[B_1] + \cdots + \mathrm{E}[B_k] = 0$$

while the variance becomes

$$\begin{aligned}\operatorname{Var}[B_1+\cdots+B_k] &= \operatorname{Var}[kB_1+(k-1)(B_2-B_1)+(k-2)(B_3-B_2)+\cdots+B_k-B_{k-1}]\\ &= k^2\operatorname{Var}[B_1]+(k-1)^2\operatorname{Var}[B_2-B_1]+(k-2)^2\operatorname{Var}[B_3-B_2]+\cdots+1\\ &= k^2+(k-1)^2+\cdots+1\end{aligned}$$

In other words

$$B_1+\cdots+B_k \sim \text{Normal}(0, 1+\cdots+k^2)$$

or, if you like,

$$B_1+\cdots+B_k \sim \text{Normal}\left(0, \frac{k(k+1)(2k+1)}{6}\right).$$

7. (a) If $\mathbf{1}$ is the vector consisting of 1's, we get that $R\mathbf{1} = \frac{1}{\lambda}Q\mathbf{1} + I\mathbf{1} = 0 + \mathbf{1}$, so all rows of R sum to 1. In addition, we need all elements of R to be non-negative. The elements of Q are non-negative except for the diagonal, which means that the off-diagonal elements of $\frac{1}{\lambda}Q + I$ are non-negative. The i'th diagonal element is $-q_i/\lambda + 1$, where q_i is the i'th diagonal element of Q. Setting $-q_i/\lambda + 1 \geq 0$ yields $\lambda \geq q_i$ for all i which is the requirement we need.

   (b) As $Q = \lambda(R - I)$ and as the diagonal matrix $e^{-t\lambda I}$ commutes with any other matrix, we get

$$e^{tQ} = e^{-t\lambda I + t\lambda R} = e^{-t\lambda I}e^{t\lambda R} = e^{-t\lambda}e^{t\lambda R}.$$

   Using the definition of the exponential matrix, we further get

$$e^{-t\lambda}e^{t\lambda R} = e^{-t\lambda}\sum_{k=0}^{\infty}\frac{(t\lambda R)^k}{k!} = e^{-t\lambda}\sum_{k=0}^{\infty}\frac{(t\lambda)^k}{k!}R^k.$$

   (c) Instead of summing to infinity in the matrix sum above, we can sum some finite N number of terms. This will usually give a better approximation to the limit than using the same number of terms in the definition

$$e^{tQ} = \sum_{k=0}^{\infty}\frac{t^k}{k!}Q^k.$$

Appendix: Some probability distributions

The Beta distribution

If $x \in [0, 1]$ has a Beta distribution with parameters with $\alpha > 0$ and $\beta > 0$ then the density is

$$\pi(x \mid \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1 - x)^{\beta-1}.$$

We write $x \mid \alpha, \beta \sim \text{Beta}(\alpha, \beta)$ and $\pi(x \mid \alpha, \beta) = \text{Beta}(x; \alpha, \beta)$.

The Beta-Binomial distribution

If $x \in \{0, 1, 2, \ldots, n\}$ has a Beta-Binomial distribution, with n a positive integer and parameters $\alpha > 0$ and $\beta > 0$, then the probability mass function is

$$\pi(x \mid n, \alpha, \beta) = \binom{n}{x} \frac{\Gamma(x + \alpha)\Gamma(n - x + \beta)\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(n + \alpha + \beta)}.$$

We write $x \mid n, \alpha, \beta \sim \text{Beta-Binomial}(n, \alpha, \beta)$ and $\pi(x \mid n, \alpha, \beta) = \text{Beta-Binomial}(x; n, \alpha, \beta)$.

The Binomial distribution

If $x \in \{0, 1, 2, \ldots, n\}$ has a Binomial distribution, with n a positive integer and $0 \le p \le 1$, then the probability mass function is

$$\pi(x \mid n, p) = \binom{n}{x} p^x (1 - p)^{n-x}.$$

We write $x \mid n, p \sim \text{Binomial}(n, p)$ and $\pi(x \mid n, p) = \text{Binomial}(x; n, p)$.

The Dirichlet distribution

If $x = (x_1, x_2, \ldots, x_n)$ has a Dirichlet distribution, with $x_i \ge 0$ and $\sum_{i=1}^n x_i = 1$ and with parameters $\alpha = (\alpha_1, \ldots, \alpha_n)$ with $\alpha_1 > 0, \ldots, \alpha_n > 0$, then the density function is

$$\pi(x \mid \alpha) = \frac{\Gamma(\alpha_1 + \alpha_2 + \cdots + \alpha_n)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\cdots\Gamma(\alpha_n)} p_1^{\alpha_1-1} p_2^{\alpha_2-1} \cdots p_n^{\alpha_n-1}.$$

We write $x \mid \alpha \sim \text{Dirichlet}(\alpha)$ and $\pi(x \mid \alpha) = \text{Dirichlet}(x; \alpha)$.

The Exponential distribution

If $x \ge 0$ has an Exponential distribution with parameter $\lambda > 0$, then the density is

$$\pi(x \mid \lambda) = \lambda \exp(-\lambda x)$$

We write $x \mid \lambda \sim \text{Exponential}(\lambda)$ and $\pi(x \mid \lambda) = \text{Exponential}(x; \lambda)$. The expectation is $1/\lambda$ and the variance is $1/\lambda^2$.

The Gamma distribution

If $x > 0$ has a Gamma distribution with parameters $\alpha > 0$ and $\beta > 0$ then the density is

$$\pi(x \mid \alpha\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x).$$

We write $x \mid \alpha,\beta \sim \text{Gamma}(\alpha,\beta)$ and $\pi(x \mid \alpha,\beta) = \text{Gamma}(x;\alpha,\beta)$.

The Geometric distribution

If $x \in \{1, 2, 3, \ldots\}$ has a Geometric distribution with parameter $p \in (0, 1)$, the probability mass function is

$$\pi(x \mid p) = p(1 - p)^{x-1}$$

We write $x \mid p \sim \text{Geometric}(p)$ and $\pi(x \mid p) = \text{Geometric}(x; p)$. The expectation is $1/p$ and the variance $(1 - p)/p^2$.

The Negative Binomial distribution

A stochastic variable x taking on as possible values any nonnegative integer has a Negative Binomial distribution if its probability mass function is given by

$$\pi(x \mid r, p) = \binom{x + r - 1}{x} \cdot (1 - p)^x p^r = \frac{\Gamma(x + r)}{\Gamma(x + 1)\Gamma(r)} (1 - p)^x p^r$$

where $r > 0$ and $p \in (0, 1)$ are parameters.

The Normal distribution

If the real x has a Normal distribution with parameters μ and σ^2, its density is given by

$$\pi(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x - \mu)^2\right).$$

We write $x \mid \mu, \sigma^2 \sim \text{Normal}(\mu, \sigma^2)$ and $\pi(x \mid \mu, \sigma^2) = \text{Normal}(x; \mu, \sigma^2)$.

The Poisson distribution

If $x \in \{0, 1, 2, \ldots\}$ has Poisson distribution with parameter $\lambda > 0$ then the probability mass function is

$$e^{-\lambda}\frac{\lambda^x}{x!}.$$

We write $x \mid \lambda \sim \text{Poisson}(\lambda)$ and $\pi(x \mid \lambda) = \text{Poisson}(x; \lambda)$. The Poisson distribution has expectation λ and variance λ.