Common Distributions

The notation $z \sim \mathrm{P}(\theta)$ means that the random variable z is sampled from (or distributed according to) the distribution P, which is parameterized by θ. When we write $\mathrm{P}(z; \theta)$ we refer to the density (assuming it exists) of P evaluated at z. Here, we provide a summary of common distributions and their parametric densities or mass functions.

Bernoulli For a binary random variable $x \in \{0, 1\}$ with $\rho \in [0, 1]$,

$$\mathrm{Bern}(x; \rho) = \rho^x (1-\rho)^{1-x}.$$

Beta For a continuous random variable $\rho \in [0, 1]$ with $a > 0$ and $b > 0$,

$$\mathrm{Beta}(\rho; a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \rho^{a-1}(1-\rho)^{b-1}.$$

The beta distribution is a conjugate prior for the Bernoulli, binomial, and negative binomial distributions.

Binomial For an integer-valued random variable $x \in \{1, \ldots, N\}$ with $N \in \mathbb{N}$ and $\rho \in [0, 1]$,

$$\mathrm{Bin}(x; N, \rho) = \binom{N}{x} \rho^x (1-\rho)^{N-x}.$$

Dirichlet For a probability vector $\boldsymbol{\pi} \in [0, 1]^K$ such that $\pi_k \geq 0$ and $\sum_k \pi_k = 1$, and parameter $\boldsymbol{\alpha} \in \mathbb{R}_+^K$,

$$\mathrm{Dir}(\boldsymbol{\pi}; \boldsymbol{\alpha}) = \frac{\Gamma\left(\sum_{k=1}^K \alpha_k\right)}{\prod_{k=1}^K \Gamma(\alpha_k)} \prod_{k=1}^K \pi_k^{\alpha_k - 1}.$$

The Dirichlet distribution is a conjugate prior to the discrete and multinomial distributions.

Exponential For a random variable $x \in \mathbb{R}_+$ with rate $\lambda \in \mathbb{R}_+$,

$$\mathrm{Exp}(x; \lambda) = \lambda e^{-\lambda x}.$$

Gamma For a nonnegative random variable $\lambda \in \mathbb{R}_+$ with shape parameter $a > 0$ and rate parameter $b > 0$,

$$\mathrm{Gamma}(\lambda; a, b) = \frac{b^a}{\Gamma(a)} \lambda^{a-1} e^{-b\lambda}.$$

The gamma distribution may also be parameterized in terms of a scale parameter, $\theta = b^{-1}$, but we do not use that parameterization in this exam.

Inverse Wishart For a random variable $\mathbf{\Sigma} \in \mathbb{R}_{\succeq 0}^{D\times D}$ (a positive semidefinite matrix) with degrees of freedom $\nu \in \mathbb{R}_+$ and scale $\mathbf{\Sigma}_0 \in \mathbb{R}_{\succeq 0}^{D\times D}$,

$$\mathrm{IW}(\mathbf{\Sigma};\, \nu_0, \mathbf{\Sigma}_0) = \frac{|\mathbf{\Sigma}_0|^{\frac{\nu_0}{2}}}{2^{\frac{\nu_0 D}{2}}\Gamma_D\left(\frac{\nu_0}{2}\right)} |\mathbf{\Sigma}|^{-\frac{\nu_0+D+1}{2}} e^{-\frac{1}{2}\mathrm{Tr}(\mathbf{\Sigma}_0\mathbf{\Sigma}^{-1})}.$$

Laplace (a.k.a. Double Exponential) For a random variable $x \in \mathbb{R}$ with rate $\lambda \in \mathbb{R}_+$,

$$\mathrm{Lap}(x;\lambda) = \frac{\lambda}{2} e^{-\lambda|x|}.$$

Multinomial For a vector of discrete counts $\boldsymbol{x} \in \mathbb{N}^K$ with $\sum_k x_k = N$ and a probability vector $\boldsymbol{\pi} \in [0,1]^K$,

$$\mathrm{Mult}(\boldsymbol{x}; N, \boldsymbol{\pi}) = \binom{N}{x_1, x_2, \dots x_K} \prod_{k=1}^{K} \pi_k^{x_k},$$

where

$$\binom{N}{x_1, x_2, \dots x_K} = \frac{N!}{x_1! \dots x_K!}.$$

Multivariate Normal For a random variable $\boldsymbol{x} \in \mathbb{R}^D$ with mean $\boldsymbol{\mu} \in \mathbb{R}^D$ and positive semidefinite covariance matrix $\mathbf{\Sigma} \in \mathbb{R}^{D\times D}$,

$$\mathcal{N}(\boldsymbol{x}; \boldsymbol{\mu}, \mathbf{\Sigma}) = (2\pi)^{-D/2} |\mathbf{\Sigma}|^{-1/2} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\top \mathbf{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\}.$$

Negative Binomial For an integer-valued random variable $x \in \mathbb{N}$ with shape parameters $\nu \in \mathbb{R}_+$ and probability $\rho \in [0,1]$,

$$\mathrm{NB}(x;\, \nu, \rho) = \binom{x+\nu-1}{x} \rho^x (1-\rho)^\nu.$$

Poisson For an integer random variable $x \in \mathbb{N}$ and a nonnegative rate parameters $\lambda \in \mathbb{R}_+$,

$$\mathrm{Po}(x;\lambda) = \frac{1}{x!}\lambda^x e^{-\lambda}.$$

Uniform For a continuous random variable $x \in \mathbb{R}$,

$$\mathrm{Unif}(x; a, b) = \begin{cases} \frac{1}{b-a} & \text{if } a < x < b, \\ 0 & \text{o.w.}. \end{cases}$$