

## Chapter 2 - Questions and Solutions

Edwin Fennell

### 2.1 Derive the mean and variance for the binomial distribution

The binomial distribution  $B(n, p)$  is defined on  $\mathbb{N}$  by the probability mass function

$$p(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & x = 0, 1, \dots, n \\ 0 & \text{else} \end{cases}$$

More abstractly, it is the distribution of total successes of  $n$  independent Bernoulli trials with success probability  $p$ . This immediately gives the expectation as

$$np$$

since the expectation of a sum of random variables is just the sum of the expectation. Since the Bernoulli trials are independent, the variance of their sum is also just the sum of their variances, so

$$np(1-p)$$

### 2.2 Derive the mean and variance for the uniform distribution

The uniform distribution  $N(a, b)$  (for  $a < b$ ) is defined on  $\mathbb{R}$  by the probability density function

$$p(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{else} \end{cases}$$

We therefore calculate the expectation as

$$\int_a^b \frac{x}{b-a} dx = \left[ \frac{x^2}{2(b-a)} \right]_a^b = \frac{a+b}{2}$$

This result is also reasonably clear from a symmetry argument.

We also have

$$\int_a^b \frac{x^3}{b-a} dx = \left[ \frac{x^3}{3(b-a)} \right]_a^b = \frac{a^2 + ab + b^2}{3}$$

which gives us the variance as

$$\frac{a^2 + ab + b^2}{3} - \frac{(a+b)^2}{4} = \frac{a^2 - 2ab + b^2}{12} = \frac{(b-a)^2}{12}$$

### 2.3 Derive the mean and covariance matrix for the multivariate normal distribution

The multivariate Gaussian  $N(\mu, \Sigma)$  is defined on  $\mathbb{R}^k$  by the probability density function

$$p(x) = \frac{1}{(2\pi)^{\frac{k}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right)$$

Then the  $i$ th element of the expectation is

$$\int x_i \frac{1}{(2\pi)^{\frac{k}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right) dx$$

$\Sigma$  is symmetric and positive definite, so by Cholesky composition we can write

$$\Sigma^{-1} = H^T H$$

If we set  $y = H(x - \mu)$  then  $dx = |H^{-1}| dy$ , and changing variables in the above integral gives us

$$\int (H_{ik}^{-1} y_k + \mu_i) \frac{1}{(2\pi)^{\frac{k}{2}}} \exp\left(-\frac{1}{2} y \cdot y\right) dy$$

We now use symmetry to discard the  $H y$  and consider all elements to obtain our expectation as

$$\mu$$

The  $(i, j)$ th element of the covariance matrix is given by

$$\int (x_i - \mu_i)(x_j - \mu_j) \frac{1}{(2\pi)^{\frac{k}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right) dx$$

Performing the previous substitution gives

$$\int H_{ia}^{-1} y_a H_{jb}^{-1} y_b \frac{1}{(2\pi)^{\frac{k}{2}}} \exp\left(-\frac{1}{2} y \cdot y\right) dy$$

This is where the real power of this substitution comes in - we could have derived the mean without it just using the symmetry around  $\mu$ , but here we would have been stuck. Observe that the substitution has made our integral separable. For any fixed  $(i, j)$  we write this integral as the sum of integrals indexed by  $a$  and  $b$  - note that the value of any constituent

integral with  $a! = b$  is necessarily 0 by symmetry in our separated integrals. Therefore we just sum over all cases where  $a = b$ , and play with indices in the result, to get the value of the integral as

$$\int (H^{-1}(H^T)^{-1})_{ij}(y \cdot y) \frac{1}{(2\pi)^{\frac{k}{2}}} \exp\left(-\frac{1}{2}y \cdot y\right) dy$$

Using results for variance of a 1D-Gaussian and separating the integrand, this simplifies to

$$(HH^T)_{ij}^{-1} = \Sigma_{ij}$$

distribution And so the covariance matrix is just

$$\Sigma$$

**2.4 Show that the mean and variance of the beta distribution with parameters  $a$  and  $b$  are**

$$\frac{a}{a+b}$$

**and**

$$\frac{ab}{(a+b)^2(a+b+1)}$$

**respectively**

The beta distribution  $Beta(a, b)$  is defined on  $[0, 1]$  by the probability density function

$$\frac{x^{a-1}(1-x)^{b-1}}{B(a, b)}$$

where

$$B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

The expectation is given by

$$\int_0^1 \frac{x^a(1-x)^{b-1}}{B(a, b)} dx = \int_0^1 \frac{\Gamma(a+b)x^a(1-x)^{b-1}}{\Gamma(a)\Gamma(b)} dx$$

We note that  $\Gamma(a+b+1) = (a+b)\Gamma(a+b)$  and  $\Gamma(a+1) = a\Gamma(a)$  so our integral becomes

$$\frac{a}{a+b} \int_0^1 \frac{\Gamma(a+b+1)x^a(1-x)^{b-1}}{\Gamma(a+1)\Gamma(b)} dx$$

The integrand is just the density function of  $Beta(a+1, b)$ , so the above expression just reduces to

$$\frac{a}{a+b}$$

Similarly, we have

$$\int_0^1 \frac{x^{a+1}(1-x)^{b-1}}{B(a,b)} dx = \int_0^1 \frac{\Gamma(a+b)x^{a+1}(1-x)^{b-1}}{\Gamma(a)\Gamma(b)} dx$$

We have  $\Gamma(a+b+2) = (a+b+1)(a+b)\Gamma(a+b)$ ,  $\Gamma(a+2) = (a+1)(a)\Gamma(a)$ , so this integral is just equal to

$$\frac{a(a+1)}{(a+b)(a+b+1)} \int_0^1 \frac{\Gamma(a+b+2)x^{a+1}(1-x)^{b-1}}{\Gamma(a+2)\Gamma(b)} dx = \frac{a(a+1)}{(a+b)(a+b+1)}$$

and therefore the variance is

$$\frac{a(a+1)}{(a+b)(a+b+1)} - \frac{a^2}{(a+b)^2} = \frac{ab}{(a+b)^2(a+b+1)}$$