Exercise Sheet 7

Robin Schmidt
Probabilistic Inference & Learning

November 27, 2018

Laplace Approximations

The original Laplace approximation

(a)

$$\widetilde{p} = r^{a-1}(1-r)^{b-a}$$
$$log\ \widetilde{p} = log(r^{a-1}(1-r)^{b-a})$$
$$= log(r^{a-1}) + log((1-r)^{b-1})$$
$$= (a-1)\ log(r) + (b-1)\ log(1-r)$$

Taking the derivatives for $g(r)$ and $\psi(r)$ gives:

$$g(r) = \frac{\partial\ log\ \widetilde{p}}{\partial r} = (a-1)\frac{1}{r} + (b-1)\frac{1}{1-r}$$
$$\psi(r) = \frac{\partial^2\ log\ \widetilde{p}}{(\partial r)^2} = -(a-1)\frac{1}{r^2} + (b-1)\frac{1}{(1-r)^2}$$

(b)

Setting $g(r) = 0$ gives $\hat{r}$ for $a, b \neq 1$:

$$\frac{a-1}{r} + \frac{b-1}{1-r} \overset{!}{=} 0$$

$$(a-1)(1-r) + (b-1)r \overset{!}{=} 0$$

$$a + ar - 1 + r + rb - r \overset{!}{=} 0$$

$$a + ar + rb - 1 \overset{!}{=} 0$$

$$ar + rb \overset{!}{=} 1 - a$$

$$r(a+b) \overset{!}{=} 1 - a$$

$$r \overset{!}{=} \frac{1-a}{a+b} \quad = \hat{r}$$

(b)

< Placeholder >

The Gamma Function

Using the definition of the Beta function from (1) and the definition of the Beta function in terms of the Gamma function we get:

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

Which yields a probability density function $f(x)$ over $[0;1]$ with:

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

Computing the mean gives:

$$m(a,b) = \mathbb{E}_{\mathcal{B}(r;a,b)}[r] = \int_0^1 r \cdot \mathcal{B}(r;a,b) \ dr$$

$$= \int_0^1 r \cdot f(r) \ dr$$

$$= \int_0^1 r(\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}r^{a-1}(1-r)^{b-1}) \ dr$$

$$= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 r^a(1-r)^{b-1} \ dr$$

Now the integral in this equation is essentially the definition we started out with, but with $\Gamma(a+1)$ instead of $\Gamma(a)$. Using this property gives the following:

$$m(a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot \frac{\Gamma(a+1)\Gamma(b)}{\Gamma((a+1)+b)}$$

$$= \frac{\Gamma(a+b)}{\Gamma(a)} \cdot \frac{\Gamma(a+1)}{\Gamma(a+b+1)}$$

$$= \frac{\Gamma(a+b)}{\Gamma(a+b+1)} \cdot \frac{\Gamma(a+1)}{\Gamma(a)}$$

Now using the property that $\Gamma(x+1) = x\Gamma(x)$ yields the desired result of:

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

For the variance we essentially use the same approach as already used for the mean. Recalling central moments for the variance of a distribution, we start out for the variance like:

$$v(a,b) = \mathbb{E}_{\mathcal{B}(r;a,b)}[r^2] - m^2(a,b) = \int_0^1 r^2 \cdot \mathcal{B}(r;a,b) \; dr - m^2(a,b)$$

$$= \int_0^1 r^2 \cdot f(r) \; dr - m^2(a,b)$$

$$= \int_0^1 r^2 \left(\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} r^{a-1}(1-r)^{b-1}\right) \; dr - \left(\frac{a}{a+b}\right)^2$$

$$= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 r^{(a+2)-1}(1-r)^{b-1} \; dr - \left(\frac{a}{a+b}\right)^2$$

Now substituting the integral with the representation in terms of the Gamma function gives:

$$v(a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot \frac{\Gamma(a+2)\Gamma(b)}{\Gamma((a+2)+b)} - \left(\frac{a}{a+b}\right)^2$$

$$= \frac{\Gamma(a+b)}{\Gamma(a+b+2)} \cdot \frac{\Gamma(a+2)}{\Gamma(a)} - \left(\frac{a}{a+b}\right)^2$$

Again using the property that $\Gamma(x + 1) = x\Gamma(x)$ gives:

$$v(a, b) = \frac{1}{(a + b)(a + b + 1)} \cdot a(a + 1) - (\frac{a}{a + b})^2$$

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

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

$$= \frac{a^3 + a^2b + a^2 + ab - (a^3 + a2b + a^2)}{(a + b)^2(a + b + 1)}$$

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

Stirling's approximation

(a)

We start by computing the logarithm of $\widetilde{p}(t|a, b)$:

$$\widetilde{p}(t|a, b) = b^a \cdot t^{a-1}e^{-bt}$$
$$log\,\widetilde{p}(t|a, b) = log(b^a \cdot t^{a-1}e^{-bt})$$
$$= log(b^a) + log(t^{a-1}) + log(e^{-bt})$$
$$= a \cdot log(b) + (a - 1) \cdot log(t) - bt$$

For the derivatives $g(t)$ and $\psi(t)$ we get:

$$g(t) = \frac{\partial log\,\widetilde{p}(t|a, b)}{\partial t} = \frac{\partial}{\partial t}(a \cdot log(b) + (a - 1) \cdot log(t) - bt)$$
$$= \frac{a - 1}{t} - b$$

$$\psi(t) = \frac{\partial^2 log\,\widetilde{p}(t|a, b)}{(\partial t)^2} = \frac{\partial^2}{(\partial t)^2}(a \cdot log(b) + (a - 1) \cdot log(t) - bt)$$
$$= -\frac{a - 1}{t^2}$$

(b)

Setting $g(t) = 0$ gives $\hat{t}$ like:

$$g(t) \overset{!}{=} 0$$

$$\frac{a-1}{t} - b \overset{!}{=} 0$$

$$\frac{a-1}{b} \overset{!}{=} t \quad = \hat{t}$$

Now substituting t for $\hat{t}$ in $\psi(t)$ gives:

$$\psi(\hat{t}) = -\frac{a-1}{(\frac{a-1}{b})^2}$$

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

$$= -\frac{b^2}{a-1}$$

(c)

< Placeholder >