

Bayesian inference, course exam 21.12.2017

1. Assume i.i.d sample Y_1, \dots, Y_n from the negative binomial distribution $\text{Neg-bin}(r_0, p)$, where r_0 is a known constant, with a beta prior for the parameter p :

$$Y_i | p \sim \text{Neg-bin}(r_0, p) \quad \text{for all } i = 1, \dots, n$$
$$p \sim \text{Beta}(\alpha, \beta).$$

Derive the posterior distribution $p(p | \mathbf{y})$ for the parameter p given all the observations $\mathbf{y} = (y_1, \dots, y_n)$.

2. Let's consider a new observation $\tilde{Y} | p \sim \text{Neg-bin}(r_0, p)$ from the same distribution as the observations of the previous exercise. Assume also that this new observation is conditionally independent from the original observations given the parameter p . Denote the parameters of the posterior distribution of the previous exercise as α_n and β_n (assume the same beta prior $\text{Beta}(\alpha, \beta)$ from the previous exercise).

Show that the posterior predictive distribution for the new observation can be written as:

$$p(\tilde{y} | \mathbf{y}) = \binom{r_0 + \tilde{y} - 1}{\tilde{y}} \frac{B(r_0 + \alpha_n, \tilde{y} + \beta_n)}{B(\alpha_n, \beta_n)}.$$

3. Assume i.i.d. observations Y_1, \dots, Y_n from the gamma distribution with both parameters α and β *unknown*:

$$Y_i | \alpha, \beta \sim \text{Gamma}(\alpha, \beta) \quad \text{for all } i = 1, \dots, n.$$

The conjugate prior for this model with hyperparameters $p, r, q, s > 0$ is of the form

$$p(\alpha, \beta | p, q, r, s) \propto \frac{\beta^{\alpha s}}{\Gamma(\alpha)^r} p^{\alpha-1} \exp(-\beta q), \quad \text{where } \alpha, \beta > 0.$$

(a) Show that given this prior, the joint posterior for the parameters α and β given all the observations has the same form:

$$p(\alpha, \beta | \mathbf{y}) \propto \frac{\beta^{\alpha s_n}}{\Gamma(\alpha)^{r_n}} p_n^{\alpha-1} \exp(-\beta q_n),$$

where the parameters of the posterior distribution are

$$p_n := p \prod_{i=1}^n y_i,$$
$$q_n := q + \sum_{i=1}^n y_i,$$
$$r_n := r + n,$$
$$s_n := s + n.$$

(b) Show that the marginal posterior for the parameter α given all the observations $\mathbf{y} = (y_1, \dots, y_n)$ is of the form:

$$p(\alpha | \mathbf{y}) \propto \frac{\Gamma(\alpha s_n + 1)}{\Gamma(\alpha)^{r_n} q_n^{\alpha s_n + 1}} p_n^{\alpha-1}, \quad \text{where } \alpha > 0.$$

Continued on the other side...

4. Explain briefly meanings of the following terms (you do not have to remember the exact mathematical definition for the last two items: 'hand-wavy' explanation is sufficient):

- (a) Noninformative prior
- (b) Probabilistic programming
- (c) Strong law of large numbers
- (d) Highest posterior density (HPD) credible interval

Density functions

The probability mass function of the random variable X following the **geometrical distribution** $\text{Geom}(p)$ is:

$$p(x | p) = p(1 - p)^x, \quad \text{when } x = 0, 1, 2, \dots$$

where $0 < p < 1$.

The density function of the random variable X following the **beta distribution** $\text{Beta}(\alpha, \beta)$ is:

$$p(x | \alpha, \beta) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1}, \quad \text{when } 0 < x < 1,$$

where $\alpha, \beta > 0$.

The probability mass function of the random variable X following the **negative binomial distribution** $\text{Neg-bin}(r, p)$ is:

$$p(x | r, p) = \binom{r+x-1}{x} p^r (1-p)^x, \quad \text{when } x = 0, 1, 2, \dots$$

where $0 < p < 1$, and $r = 1, 2, \dots$

The density function of the random variable X following the **beta distribution** $\text{Beta}(\alpha, \beta)$ is:

$$p(x | \alpha, \beta) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1}, \quad \text{when } 0 < x < 1,$$

where $\alpha, \beta > 0$.

The density of the random variable X following the **gamma distribution** $\text{Gamma}(\alpha, \beta)$ is:

$$p(x | \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x), \quad \text{when } x > 0,$$

where $\alpha, \beta > 0$.

Integrals

- Gaussian integral:

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}.$$

- Gamma integral:

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt.$$

- Beta function:

$$B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} = \int_0^1 t^{a-1} (1-t)^{b-1} dt.$$