EST-25134: Aprendizaje Estadístico – ITAM

Primavera 2017

Sesión 01

Funciones de distribución

Traten de identificar el kernel y constante de normalización en las siguientes distribuciones:

- *Distribución Bernoulli.*- Para X v.a. con soporte en $\mathcal{X} = \{0, 1\}$, la función de masas de probabilidades es,
$$f_X(x|\alpha, \beta) = \alpha^x (1-\alpha)^{1-x} \mathbf{1}_{\{0,1\}}(x), \tag{1}$$
con $0 < \alpha < 1$.

- *Distribución beta.*- Para X v.a. con soporte en $\mathcal{X} = (0, 1)$, la función de densidad,
$$f_X(x|\alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} \mathbf{1}_{(0,1)}(x), \tag{2}$$
con $\alpha, \beta > 0$.

- *Distribución multinomial.*- Para X variable aleatoria con soporte en $\mathcal{X} = \{0, 1\}^K$, con K finito y conocido, tenemos que la función de masas de probabilidades es,
$$f_X(x|\alpha_1, \ldots, \alpha_K) = \prod_{k=1}^{K} \alpha_k^{x_k} \mathbf{1}_{\{0,1\}^K}(x), \tag{3}$$
donde $\alpha_k \geq 0$ y $\sum_{k=1}^{K} = 1$.

- *Distribución Dirichlet.*- Para X v.a. con soporte en $\mathcal{X} = \mathcal{S}_K$ (simplex de dimensión $(K-1)$), de tiene que la función de densidad es,
$$f_X(x|\alpha_1, \ldots, \alpha_K) = \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_K)} \prod_{k=1}^{K} x_k^{\alpha_k} \mathbf{1}_{\mathcal{S}_K}(x), \tag{4}$$
donde $\alpha_0 = \sum_{k=1}^{K} \alpha_k$, con $\alpha_k > 0$.

Distribución gamma-inversa

En el contexto bayesiano del modelo de regresión, pensando en σ^2 como la varianza condicional de y dado $\boldsymbol{x}$ (i.e. modelo homoscedástico), la distribución inicial para σ^2 es gamma-inversa, si su función de densidad es de la forma,

$$\begin{aligned} \pi(\sigma^2) &= \frac{b^a}{\Gamma(a)} \left(\frac{1}{\sigma^2}\right)^{a+1} \exp\left\{-b/\sigma^2\right\} \mathbf{1}_{(0,\infty)}(\sigma^2) & (5) \\ &\propto \exp\left\{-b/\sigma^2\right\} \mathbf{1}_{(0,\infty)}(\sigma^2), & (6) \end{aligned}$$

donde $a, b > 0$.

Distribución normal-gamma-inversa

De nuevo, en el contexto del modelo de regresión, donde

$$y|\boldsymbol{x}, \beta, \sigma^2 \sim N(y|\boldsymbol{x}'\beta, \sigma^2), \tag{7}$$

donde $\boldsymbol{x} = (x_1, \ldots, x_p)$ y $\beta = (\beta_1, \ldots, \beta_p)$, se considera el caso donde (β, σ^2) son parámetros aleatorios. Es común entonces encontrar que la distribución conjunta para estos parámetros es normal-gamma-inversa, i.e. $\beta|\sigma^2$ se distribuye normal (o gaussiana) y σ^2 se distribuye marginalmente gamma inversa.

La densidad de la distribución normal-gamma inversa es de la forma,

$$\begin{aligned}
\pi(\beta, \sigma^2) &= \pi(\beta|\sigma^2)\pi(\sigma^2) \\
&= N_p(\beta|m, \sigma^2 S) GaInv(\sigma^2|a, b) \\
&= \frac{b^a}{(2\pi)^{p/2}|S|^{1/2}\Gamma(a)} \left(\frac{1}{\sigma^2}\right)^{a+1} \exp\left\{-\frac{1}{\sigma^2}\left(b + 1/2(\beta - m)' S^{-1}(\beta - m)\right)\right\} \mathbf{1}_{\Re^p \times \Re_+}(\beta, \sigma^2) \\
&\propto \left(\frac{1}{\sigma^2}\right)^{a+1} \exp\left\{-\frac{1}{\sigma^2}\left(b + 1/2(\beta - m)' S^{-1}(\beta - m)\right)\right\} \mathbf{1}_{\Re^p \times \Re_+}(\beta, \sigma^2), \qquad (8)
\end{aligned}$$

donde $a, b > 0$, m es un vector p-dimensional y S es una matriz positivo definida simétrica.

Distribución t-Student multivariada

De la distribución normal-gamma-inversa se sigue que la distribución marginal para β (integrando σ^2) es t-Student. El cálculo se obtiene como sigue,

$$\begin{aligned}
\pi(\beta) &= \int_{\Re_+} \pi(\beta, \sigma^2) d\sigma^2 \\
&= \frac{b^a}{(2\pi)^{p/2}|S|^{1/2}\Gamma(a)} \int_{\Re_+} \left(\frac{1}{\sigma^2}\right)^{a+1} \exp\left\{-\frac{1}{\sigma^2}\left(b + 1/2(\beta - m)' S^{-1}(\beta - m)\right)\right\} d\sigma^2 \\
&= \frac{b^a \Gamma(a + p/2)}{(2\pi)^{p/2}|S|^{1/2}\Gamma(a)} \left(b + 1/2(\beta - m)' S(\beta - m)\right)^{-(a+p/2)} \mathbf{1}_{\Re^p}(\beta). \qquad (9)
\end{aligned}$$

Esta densidad corresponde a la de la distribución t-Student con $2a$ grados de libertad.