Chapter 3

Examples of Mass Functions and Densities

Exponential and Location/Scale Families

Outline

Introduction

Let the parameter space Θ be a non-empty open subset of $\mathbb{R}^k$. Thus, $\eta \in \Theta$ can be vector valued. A family of continuous (discrete) random variables is called an exponential family if the probability density functions (probability mass functions) can be expressed in the form

$$f_X(x|\eta) = h(x) \exp\left(\sum_{i=1}^{k} \eta_i t_i(x) - A(\eta)\right), \quad x \in S.$$

Thus, S, the sate space, is common for all the $f_X(x|\eta)$.

$h(x)$ is a non-negative function.
The $t_i(x)$ are real-valued functions on the state space.
$A(\eta)$ is called the cumulant function..

Under this expression for an exponential family, η is called the natural parameter.

Examples

Example. Fix a value for r. For X a negative binomial random variable,

$$f_X(x|p) = \binom{x+r-1}{x}(1-p)^r p^x.$$

Let $\eta = \ln p$, then we can write this expression

$$f_X(x|\eta) = \binom{x+r-1}{x}(1-e^\eta)^r e^{\eta x}.$$

Thus,

$$h(x) = \binom{x+r-1}{x}, \quad A(\eta) = -r\ln(1-e^\eta) \quad t(x) = x.$$

Examples

Example. For X a gamma random variable,

$$f_X(x|\alpha,\beta) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} & \text{for } 0 \le x, \\ 0 & \text{otherwise.} \end{cases}$$

We rewrite the density to obtain the form of an exponential family.

$$f_X(x|\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} = x^{-1} \frac{\beta^\alpha}{\Gamma(\alpha)} \exp(\alpha \ln x - \beta x)$$

Thus, $\eta_1 = \alpha$ and $\eta_2 = -\beta$,

$$h(x) = x^{-1}, \quad A(\eta_1,\eta_2) = -\ln \frac{(-\eta_2)^{\eta_1}}{\Gamma(\eta_1)}, \quad t_1(x) = \ln x, \quad t_2(x) = x$$

, Then for $x \ge 0$

$$f_X(x|\eta) = x^{-1} \exp\left((\eta_1 \ln x + \eta_2 x) + \ln \Gamma(\eta_1) - \eta_1 \ln(-\eta_2)\right).$$

Examples

Example. For X a normal random variable,

$$f_X(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp -\frac{(x-\mu)^2}{2\sigma^2} = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-x^2}{2\sigma^2} + \frac{\mu x}{\sigma^2} - \frac{\mu^2}{2\sigma^2}\right)$$

Thus, $\eta_1 = \mu/\sigma^2$ and $\eta_2 = -1/(2\sigma^2)$. Here,

$$t_1(x) = x, \quad t_2(x) = x^2, \quad h(x) = \frac{1}{\sqrt{2\pi}}, \quad A(\eta) = -\frac{1}{2}\ln(-2\eta_2) - \frac{\eta_1^2}{4\eta_2}.$$

Expected Values

Note that for continuous random variables from an exponential family

$$1 = \int f_X(x|\eta)\ dx = \int h(x) \exp\left(\sum_{i=1}^{k} \eta_i t_i(x) - A(\eta)\right)\ dx.$$

$$\exp A(\eta) = \int h(x) \exp\left(\sum_{i=1}^{k} \eta_i t_i(x)\right)\ dx.$$

Take the partial derivative with respect to η_i to obtain

$$e^{A(\eta)} \frac{\partial A(\eta)}{\partial \eta_i} = \int t_i(x) h(x) \exp\left(\sum_{i=1}^{k} \eta_i t_i(x)\right)\ dx,$$

$$\frac{\partial A(\eta)}{\partial \eta_i} = \int t_i(x) h(x) \exp\left(\sum_{i=1}^{k} \eta_i t_i(x) - A(\eta)\right)\ dx = \int t_i(x) f_X(x|\eta)\ dx = Et_i(X)$$

Expected Values

Replace the integral by a sum to obtain the same identity for discrete random variables.

$$Et_i(X) = \frac{\partial A(\eta)}{\partial \eta_i}.$$

For X a negative binomial random variable, $\eta = \ln p$

$$A(\eta) = -r\ln(1-e^{\eta}), \quad t(x) = x.$$

Thus,

$$EX = \frac{\partial A(\eta)}{\partial \eta_i} = \frac{re^{\eta}}{1-e^{\eta}} = \frac{rp}{1-p}$$

Expected Values

For X a gamma random variable, $\alpha = \eta_1$, $\beta = -\eta_2$, $t_1(x) = \ln x$, $t_2(x) = x$

$$A(\eta) = -\eta_1 \ln(-\eta_2) + \ln \Gamma(\eta_1).$$

Then

$$\begin{aligned} E[\ln X] &= \frac{\partial}{\partial \eta_1}(-\eta_1 \ln(-\eta_2) + \ln \Gamma(\eta_1)) = -\ln(-\eta_2) + \frac{d}{d\eta_1} \ln \Gamma(\eta_1) \\ &= -\ln(\beta) + \frac{d}{d\alpha} \ln \Gamma(\alpha), \\ EX &= \frac{\partial}{\partial \eta_2}(-\eta_1 \ln(-\eta_2) + \ln \Gamma(\eta_1)) = -\frac{\eta_1}{\eta_2} = \frac{\alpha}{\beta} \end{aligned}$$

$\psi(\alpha) = \frac{d}{d\alpha} \ln \Gamma(\alpha)$ is called the digamma function. The R command is `digamma`.

Variances

For a one parameter family continuous random variables from an exponential family

$$1 = \int f_X(x|\eta)\ dx = \int h(x) \exp\left(\eta t(x) - A(\eta)\right)\ dx.$$

$$\exp A(\eta) = \int h(x) \exp\left(\eta t(x)\right)\ dx.$$

Take k derivatives with respect to η to obtain

$$\begin{aligned} \frac{d^k}{d\eta^k} e^{A(\eta)} &= \int t(x)^k h(x) \exp\left(\eta t(x)\right)\ dx, \\ e^{-A(\eta)} \frac{d^k}{d\eta^k} e^{A(\eta)} &= \int t(x)^k h(x) \exp\left(\eta t(x) - A(\eta)\right)\ dx = Et_i(X)^k \end{aligned}$$

Variances

Let's consider more carefully the case $k = 2$. As before,

$$e^{-A(\eta)}\frac{d}{d\eta}e^{A(\eta)} = A'(\eta) = Et(X).$$

For the second derivative,

$$\frac{d^2}{d^2\eta}e^{A(\eta)} = \frac{d}{d\eta}\left(A'(\eta)e^{A(\eta)}\right) = \left(A''(\eta) + A'(\eta)^2\right)e^{A(\eta)}$$

$$Et(X)^2 = e^{-A(\eta)}\frac{d^2}{d^2\eta}e^{A(\eta)} = A''(\eta) + A'(\eta)^2$$

$$Et(X)^2 = A''(\eta) + (Et(X))^2$$

$$A''(\eta) = Et(X)^2 - (Et(X))^2 = \mathrm{Var}(t(X))$$

Variances

For X a negative binomial random variable, $\eta = \ln p$,

$$A(\eta) = -r\ln(1-e^\eta) \quad t(x) = x.$$

and

$$\begin{aligned} \mathrm{Var}(x) &= A''(\eta) = \frac{d}{d\eta}\frac{re^\eta}{1-e^\eta} = \frac{r(1-e^\eta)e^\eta - e^\eta(-e^\eta))}{(1-e^\eta)^2} \\ &= \frac{re^\eta}{(1-e^\eta)^2} = \frac{rp}{(1-p)^2} \end{aligned}$$

For X a Poisson random variable,

$$f_X(x|\lambda) = \frac{\lambda^x}{x!}\exp(-\lambda) = \frac{1}{x!}\exp(x\ln\lambda - \lambda)$$

$$\eta = \ln\lambda, \qquad A(\eta) = e^\eta.$$

Thus, $\mathrm{Var}(X) = A''(\eta) = e^\eta = \lambda$.

Variances

Example. For X a normal random variable,

$$t_1(x) = x, \quad \eta_1 = \mu/\sigma^2, \quad \eta_2 = -1/(2\sigma^2) \quad A(\eta) = \frac{1}{2}\ln(-2\eta_2) - \frac{\eta_1^2}{4\eta_2}.$$

$$EX = \frac{\partial A(\eta)}{\partial \eta_1} = \frac{-\eta_1}{2\eta_2} = \frac{\mu/\sigma^2}{2/(2\sigma^2)} = \mu$$

$$\mathrm{Var}(X) = \frac{\partial^2 A(\eta)}{\partial \eta_1} = \frac{-1}{2\eta_2} = \frac{1}{2/(2\sigma^2)} = \sigma^2$$

The Logistic Distribution

Let C_0 and C_1 be a partition of the probability space Ω with $\pi = P(C_0)$ and consider an exponential family.

$$f_X(x|\eta) = h(x)\exp\left(\eta t(x) - A(\eta)\right).$$

Assume that $t(x)$ is one-to-one. If $\omega \in C_i$, then X is distributed according to $f_X(x|\eta_i), i = 0, 1$. Thus, X is a mixture,

$$f_X(x) = f_X(x|\eta_0)\pi_0 + f_X(x|\eta_1)(1-\pi_0).$$

For example, C_0 and C_1 could represent a control and treatment group and the two distributions could be derived from models of the values for x (reading scores under different reading programs, blood pressure under two medications, grain size under two variants of corn, wave height under two ocean weather conditions, etc.)

The Logistic Distribution

To see how the logistic distribution arises, we will use Bayes formula to determine the probability that an observation comes from the control given the value of X.

$$
\begin{aligned}
P(C_0|X=x) &= \frac{P\{X=x|C_0\}P(C_0)}{P\{X=x\}} = \frac{\pi_0 f_X(x|\eta_0)}{\pi_0 f_X(x|\eta_0) + (1-\pi_0) f_X(x|\eta_1)} \\
&= \frac{1}{1+((1-\pi_0)/\pi_0)\cdot f_X(x|\eta_1)/f_X(x|\eta_0)} \\
&= \frac{1}{1+((1-\pi_0)/\pi_0)\exp(-A(\eta_1)+A(\eta_0))\cdot\exp((\eta_1-\eta_0)T(x))} \\
P(C_0|T(X)=t) &= \frac{1}{1+\exp(\alpha+\beta t)} = F_T(t|\alpha,\beta), \quad \text{the logistic distribution}
\end{aligned}
$$

Here, $\alpha = \ln((1-\pi_0)/\pi_0) - A(\eta_1) + A(\eta_0)$ and $\beta = \eta_1 - \eta_0$.

Location/Scale Families

Let X be a continuous (discrete) random variable with density (mass) function $f_X(x)$. Let

$$Y = \tau X + \gamma \quad \tau > 0, \gamma \in \mathbb{R}.$$

Then Y has density (mass) function,

$$f_Y(y|\gamma,\tau) = \frac{1}{\tau} f_X((y-\gamma)/\tau), \qquad f_Y(y|\gamma,\tau) = f_X((y-\gamma)/\tau).$$

Such a two parameter family of density (mass) functions is called a location/scale family.

- γ is the location parameter. If X has mean 0, then γ is the mean of Y.
- τ is the scale parameter. If X has standard deviation 1, then τ is the standard deviation of Y. The case $\gamma = 0$ is called a scale family.

Examples

- Uniform (both discrete and continuous), normal, and logistic random variables are examples of location-scale families.
- t is a location family
- Exponential, gamma (β), and Pareto (α) are example of scale families.

The standard Cauchy distribution for X, with density

$$f_X(x) = \frac{1}{\pi}\frac{1}{1+x^2}, \quad x \in \mathbb{R},$$

can be turned into a location-scale family.

Examples

Exercise. A Cauchy random variable does not have a mean.

Answer. The improper integral

$$\int_0^\infty \frac{1}{\pi}\frac{x}{1+x^2}\ dx$$

does not converge.

NB. The Cauchy family of distributions is *not* an exponential family