University of California, Los Angeles
Department of Statistics

Statistics 100B **Instructor: Nicolas Christou**

A note on expectation and independence and exponential families

In this note we will discuss results on expectation of functions of random variables for one and two random variables and how independence is established between two random variables. Finally we will discuss exponential families and how to identify an exponential family and how they are used to find the mean and variance of a random random variable.

1. Let X be a continuous random variable. Then $E[X] = \int_x xf(x)dx$.

2. Suppose we want to find the expectation of a function of X. Let $Y = g(X)$. Show that $E[g(X)] = \int_x g(x)f(x)dx$.

   One way to compute $E[Y]$ is to find the pdf of Y and then
   $E[Y] =$.

   So let's find $f(y)$. Use the method of cdf. Begin with the cdf of Y.

   $$\begin{aligned} F_Y(y) &= P[Y \leq y] = P[g(X) \leq y] = P[X \leq w(y)] \\ F_Y(y) &= F_X[w(y)] \end{aligned}$$

   Take derivative on both sides w.r.t. y to get
   $f_Y(y) =$.

   Back to the proof:
   Let $I = \int_x g(x)f(x)dx$. We will show that this is equal to $E[g(X)]$.

   Let $y = g(x)$ and solve for x. We get $x = w(y)$. Complete the following:
   $\frac{dx}{dy} =$

   Transform the integral I in terms of y: $I = \int_y$

   What do you observe?

   Example:
   Let $X \sim exp(1)$, then $f(x) =$. Find $E[X^3]$.

Answer:
$E[X^3] = \int_0^\infty \qquad\qquad\qquad dx.$

To evaluate this integral we can use the kernel function of the gamma distribution. Let $X \sim \Gamma(\alpha, \beta), \alpha > 0.\beta > 0, x > 0$. Then $f(x) = \frac{x^{\alpha-1}e^{-\frac{x}{\beta}}}{\Gamma(\alpha)\beta^\alpha}$. The part of the pdf $x^{\alpha-1}e^{-\frac{x}{\beta}}$ it is called the kernel function.

We also need a note on the gamma function:
Definition:
$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx.$

Properties:

$$\begin{aligned}
\Gamma(\alpha) &= (\alpha-1)\Gamma(\alpha-1) \\
\Gamma(\alpha+1) &= \alpha\Gamma(\alpha) \\
\Gamma(\alpha+2) &= \alpha(\alpha+1)\Gamma(\alpha) \\
\Gamma(\alpha) &= (\alpha-1)! \text{ (if } \alpha \text{ is an integer)}
\end{aligned}$$

Use the notes above to evaluate $E[X^3] = \int_0^\infty x^3 e^{-x}dx$.

A note on independence.
Let X, Y be random variables with joint pdf $f(x,y)$. Then X, Y are independent if

$$\begin{aligned} f(x,y) &= f(x)f(y) \\ \text{(joint pdf)} &= \text{(marginal pdf of } x\text{)(marginal pdf of } y\text{)} \end{aligned}$$

Note: To find the marginal pdf use
$f(x) = \int_y f(x,y)dy$
$f(y) = \int_x f(x,y)dx$

Theorem:
Let X, Y be independent random variables. Then $E[XY] = [EX][EY]$.
Proof: XY is a function of X and Y. Therefore, using the expectation of a function of x and y
$E[g(x,y)] = \int_x \int_y g(x,y)f(x,y)dxdy$ we get:

$$\begin{aligned} E[XY] &= \int_x \int_y xyf(x,y)dxdy \quad \text{but } X, Y \text{ are independent} \\ &= \\ &= \\ &= \end{aligned}$$

Theorem:
Let X, Y be independent random variables and let $g(x)$ and $h(y)$ be functions of x and y alone respectively. Then $E[g(X)h(Y)] = [Eg(X)]E[h(Y)]$.

Proof: Use the proof of the previous theorem.

Theorem:
Let $g(x)$ be a function of X alone and $h(y)$ be a function of Y alone. Then X, Y are independent iff $f(x, y) = g(x)h(y)$.

Proof:
Let $c = \int_{-\infty}^{\infty} g(x)dx$ and $d = \int_{-\infty}^{\infty} h(y)dy$. Show that $cd = 1$.

Now find the marginal of X and the marginal of Y. Remember that to find the marginal of X we integrate the joint pdf w.r.t. y. Here the joint pdf is given by $f(x, y) = g(x)h(y)$.

Exponential families

A probability density function or probability mass function is called an exponential family if it can be expressed as

$$f(x|\boldsymbol{\theta}) = h(x)c(\boldsymbol{\theta})exp\left(\sum_{i=1}^{k} w_i(\boldsymbol{\theta})t_i(x)\right).$$

Note: $h(x), t_1(x), \ldots, t_k(x)$ do not depend on $\boldsymbol{\theta}$ and $c(\boldsymbol{\theta})$ does not depend of x.

Example:
Consider $X \sim b(n, p)$ with n fixed. Show that $p(x) = \binom{n}{x}p^x(1-p)^{n-x}$ can be expressed in the exponential family form.

$$\begin{aligned}
p(x) &= \binom{n}{x}p^x(1-p)^{n-x} \\
&= \binom{n}{x}\left(\frac{p}{1-p}\right)^x (1-p)^n \\
&= \binom{n}{x}(1-p)^n e^{log(\frac{p}{1-p})^x} \\
&= \binom{n}{x}(1-p)^n e^{xlog(\frac{p}{1-p})}
\end{aligned}$$

Therefore this pmf is an exponential family with
$h(x) = \binom{n}{x}, c(p) = (1-p)^n, t_1(x) = x, w_1(p) = log\frac{p}{1-p}$.

Example:
Let $X \sim$ Poisson(λ). Show that $p(x) = \frac{\lambda^x e^{-\lambda}}{x!}$ is an exponential family.

Theorem:
Suppose a random variable X has a pdf or pmf that can be expressed in the form of exponential family. Then,

$$\text{(a)} \qquad E\left(\sum_{i=1}^{k} \frac{\partial w_i(\boldsymbol{\theta})}{\partial \theta_j} t_i(x)\right) = -\frac{\partial}{\partial \theta_j} logc(\boldsymbol{\theta}).$$

and

$$\text{(b)} \qquad var\left(\sum_{i=1}^{k} \frac{\partial w_i(\boldsymbol{\theta})}{\partial \theta_j} t_i(x)\right) = -\frac{\partial^2}{\partial \theta_j^2} logc(\boldsymbol{\theta}) - E\left(\sum_{i=1}^{k} \frac{\partial^2 w_i(\boldsymbol{\theta})}{\partial \theta_j^2} t_i(x)\right).$$

Note: Here log is the natural logarithm.

Proof of (a):

$$\begin{aligned}
\int_x f(x|\boldsymbol{\theta})dx &= 1 \\
\int_x h(x)c(\boldsymbol{\theta})exp\left(\sum_{i=1}^{k} w_i(\boldsymbol{\theta})t_i(x)\right)dx &= 1
\end{aligned}$$

Differentiate both sides w.r.t. θ_j:

$$\begin{aligned}
&\int_x h(x)\frac{\partial c(\boldsymbol{\theta})}{\partial \theta_j} exp\left(\sum_{i=1}^{k} w_i(\boldsymbol{\theta})t_i(x)\right)dx \\
+ &\int_x h(x)c(\boldsymbol{\theta})\sum_{i=1}^{k}\frac{\partial w_i(\boldsymbol{\theta})}{\partial \theta_j}t_i(x)exp\left(\sum_{i=1}^{k} w_i(\boldsymbol{\theta})t_i(x)\right)dx = 0
\end{aligned}$$

Multiply the first integral by $\frac{c(\boldsymbol{\theta})}{c(\boldsymbol{\theta})}$ and note that $\frac{\partial logc(\boldsymbol{\theta})}{\partial \theta_j} = \frac{\partial c(\boldsymbol{\theta})}{\partial \theta_j}\frac{1}{c(\boldsymbol{\theta})}$.

$$\begin{aligned}
&\int_x h(x)\frac{\partial c(\boldsymbol{\theta})}{\partial \theta_j} exp\left(\sum_{i=1}^{k} w_i(\boldsymbol{\theta})t_i(x)\right)\frac{c(\boldsymbol{\theta})}{c(\boldsymbol{\theta})}dx \\
+ &\int_x h(x)c(\boldsymbol{\theta})\sum_{i=1}^{k}\frac{\partial w_i(\boldsymbol{\theta})}{\partial \theta_j}t_i(x)exp\left(\sum_{i=1}^{k} w_i(\boldsymbol{\theta})t_i(x)\right)dx = 0
\end{aligned}$$

After rearranging we get

$$\begin{aligned}
&\int_x \sum_{i=1}^{k}\frac{\partial w_i(\boldsymbol{\theta})}{\partial \theta_j}t_i(x)h(x)c(\boldsymbol{\theta})exp\left(\sum_{i=1}^{k} w_i(\boldsymbol{\theta})t_i(x)\right)dx = \\
&\qquad -\frac{\partial logc(\boldsymbol{\theta})}{\partial \theta_j}\int_x h(x)c(\boldsymbol{\theta})exp\left(\sum_{i=1}^{k} w_i(\boldsymbol{\theta})t_i(x)\right)dx
\end{aligned}$$

Or

$$E\left(\sum_{i=1}^{k}\frac{\partial w_i(\boldsymbol{\theta})}{\partial \theta_j}t_i(x)\right) = -\frac{\partial}{\partial \theta_j}logc(\boldsymbol{\theta}).$$

To prove statement (b) of the theorem differentiate a second time and rearrange.

Example:
Let $X \sim$ Poisson(λ). Use the theorem above to show that $E[X] = \lambda$ and var$[X] = \lambda$.