CS6190: Probabilistic Modeling Homework 0

Gopal Menon

20 January, 2018

1. Let X and Y be continuous real-valued random variables. Prove the following:

   (a) $E\left[E\left[X|Y\right]\right] = E\left[X\right]$

$$\begin{aligned}
E\left[E\left[X|Y\right]\right] &= \int_{-\infty}^{\infty} E\left[X|Y=y\right] p_Y(y)dy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\ p_{X|Y=y}(x)dx\ p_Y(y)dy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\ p_{X|Y=y}(x)p_Y(y)dxdy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\ p_{XY}(x,y)dxdy \\
&= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} p_{XY}(x,y)dy\ dx \\
&= \int_{-\infty}^{\infty} x\ p_X(x)dx \\
&= E\left[X\right]
\end{aligned}$$

   Here $p_X(x)$, $p_Y(y)$ and $p_{X|Y=y}(x)$ are the probability density functions for X, Y, and the conditional probability of X given the value of Y. $p_{XY}(x,y)$ is the joint probability distribution function of X and Y. I needed to look up the proof [1] as I was not aware that the first step shown above could be done and due to this, I was getting an extra y in the final result. The result I got before I looked up the proof was $yE\left[X\right]$.

   (b) $Var(X) = E\left[Var(X|Y)\right] + Var(E\left[X|Y\right])$

$$\begin{aligned}
E\left[Var(X|Y)\right] + Var(E\left[X|Y\right]) &= E\left[E\left[X^2|Y\right] - E\left[X|Y\right]E\left[X|Y\right]\right] \\
&+ E\left[E\left[X|Y\right]E\left[X|Y\right]\right] - E\left[E\left[X|Y\right]\right]E\left[E\left[X|Y\right]\right] \\
&= E\left[E\left[X^2|Y\right]\right] - E\left[E\left[X|Y\right]E\left[X|Y\right]\right] \\
&+ E\left[E\left[X|Y\right]E\left[X|Y\right]\right] - E\left[X\right]E\left[X\right] \\
&= E\left[X^2\right] - E\left[X\right]E\left[X\right] \\
&= Var(X)
\end{aligned}$$

   Before I worked out the solution above, I was using the definition $Var(X) = E\left[X^2\right] - \mu_X^2$ and that did not help. After I realized that $\mu_X = E\left[X\right]$, the solution fell in place.

2. Consider random variables X and Y with joint pdf

$$p(x,y) = \begin{cases} x+y, & \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, & \text{otherwise} \end{cases}$$

(a) What is the marginal pdf $p(x)$?

$$\begin{aligned}
p(x) &= \int_{-\infty}^{\infty} p(x,y)dy \\
&= \begin{cases} \int_0^1 (x+y)dy, \text{for } 0 \le x \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \left(xy + \frac{1}{2}y^2\right)\Big|_0^1, \text{for } 0 \le x \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} x + \frac{1}{2}, \text{for } 0 \le x \le 1 \\ 0, \text{otherwise} \end{cases}
\end{aligned}$$

(b) What is the conditional pdf $p(y|x)$?

$$\begin{aligned}
p(y|x) &= \frac{p(x,y)}{p(x)} \\
&= \begin{cases} \frac{x+y}{x+\frac{1}{2}}, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases}
\end{aligned}$$

(c) What is the conditional expectation $E[Y|X]$?

$$\begin{aligned}
E[Y|X] &= \int_{-\infty}^{\infty} p_{Y|X}(y)dy \\
&= \begin{cases} \int_0^1 \frac{x+y}{x+\frac{1}{2}}dy, \text{for } 0 \le x \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \frac{xy+\frac{1}{2}y^2}{x+\frac{1}{2}}\Big|_0^1, \text{for } 0 \le x \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \frac{x+\frac{1}{2}}{x+\frac{1}{2}}, \text{for } 0 \le x \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} 1, \text{for } 0 \le x \le 1 \\ 0, \text{otherwise} \end{cases}
\end{aligned}$$

(d) What is the covariance $Cov(X,Y)$?

$$Cov(X,Y) = E[XY] - E[X]E[Y]$$

$$\begin{aligned}
E[XY] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyp(x,y)dxdy \\
&= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} yp(x,y)dydx
\end{aligned}$$

$$
\begin{aligned}
\int_{-\infty}^{\infty} yp(x,y)dy &= \begin{cases} \displaystyle\int_0^1 y(x+y)dy, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \displaystyle\int_0^1 (xy+y^2)dy, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \displaystyle\frac{1}{2}xy^2 + \frac{1}{3}y^3\Big|_0^1, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \displaystyle\frac{1}{2}x + \frac{1}{3}, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
\int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} yp(x,y)dydx &= \begin{cases} \displaystyle\int_0^1 x\left(\frac{1}{2}x + \frac{1}{3}\right)dx, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \displaystyle\int_0^1 \left(\frac{1}{2}x^2 + \frac{1}{3}x\right)dx, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \displaystyle\left(\frac{1}{2}\frac{1}{3}x^3 + \frac{1}{3}\frac{1}{2}x^2\right)\Big|_0^1, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \displaystyle\left(\frac{1}{6} + \frac{1}{6}\right), \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \displaystyle\frac{1}{3}, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= E\left[XY\right] \\
E\left[X\right] &= \int_{-\infty}^{\infty} xp(x)dx \\
&= \begin{cases} \displaystyle\int_{-\infty}^{\infty} x\left(x + \frac{1}{2}\right)dx, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \displaystyle\int_{-\infty}^{\infty} \left(x^2 + \frac{1}{2}x\right)dx, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \displaystyle\left(\frac{1}{3}x^3 + \frac{1}{2}\frac{1}{2}x^2\right)\Big|_0^1, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \displaystyle\frac{1}{3} + \frac{1}{4}, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \displaystyle\frac{7}{12}, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
p(y) &= \int_{-\infty}^{\infty} p(x,y)dx \\
&= \begin{cases} \int_0^1 (x+y)dx, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \left(\frac{1}{2}x^2 + yx\right)\Big|_0^1, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} y + \frac{1}{2}, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
E[Y] &= \int_{-\infty}^{\infty} yp(y)dy \\
&= \begin{cases} \int_{-\infty}^{\infty} y\left(y + \frac{1}{2}\right)dy, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \int_{-\infty}^{\infty} \left(y^2 + \frac{1}{2}y\right)dy, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \left(\frac{1}{3}y^3 + \frac{1}{2}\frac{1}{2}y^2\right)\Big|_0^1, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \frac{1}{3} + \frac{1}{4}, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \frac{7}{12}, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
Cov(X,Y) &= E[XY] - E[X]E[Y] \\
&= \begin{cases} \frac{1}{3} - \frac{7}{12}\frac{7}{12}, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \frac{48}{144} - \frac{49}{144}, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} -\frac{1}{144}, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases}
\end{aligned}
$$

3. Let $X \sim Exp(\lambda)$, i.e., the exponential distribution with pdf $p(x) = \lambda exp(-\lambda x)$. Let $Y = \sqrt{X}$.
   (a) What is the density function $p(y)$?

$$
\begin{aligned}
Y &= \sqrt{X} \\
f(x) &= \sqrt{x} \\
f^{-1}(y) &= y^2 \\
p(y) &= \left|\frac{d}{dy}(f^{-1}(y))\right| p(f^{-1}(y)) \\
&= \left|\frac{d}{dy}y^2\right| p(y^2) \\
&= \begin{cases} 2y\lambda exp(-\lambda y^2), \text{for } y \geq 0 \\ 0, \text{otherwise} \end{cases}
\end{aligned}
$$

(b) What is the cdf, $F(y) = P(Y \leq y)$? Verify that $F(0) = 0$ and $F(\infty) = 1$.

The cdf can be found by integrating the pdf.

$$
F(y) = P(Y \leq y) = \int_0^y 2y\lambda exp(-\lambda y^2)dy
$$

$$
\text{Substituting } u = -\lambda y^2, \text{ we get } dy = -\frac{1}{2\lambda y}du
$$

$$
\begin{aligned}
\int 2y\lambda exp(-\lambda y^2)dy &= \int 2y\lambda exp(u) \times -\frac{1}{2\lambda y}du \\
&= -\int exp(u)du \\
&= -exp(u) \\
&= -exp(-\lambda y^2) \\
F(y) &= -exp(-\lambda y^2)\Big|_0^y \\
&= -exp(-\lambda y^2) - (-exp(-\lambda 0^2)) \\
&= -exp(-\lambda y^2) + 1 \\
&= 1 - exp(-\lambda y^2) \\
F(0) &= 1 - exp(-\lambda 0^2) \\
&= 1 - 1 \\
&= 0 \\
F(\infty) &= 1 - exp(-\infty) \\
&= 1 - 0 \\
&= 1
\end{aligned}
$$

(c) What is the quantile function F^{-1}?

If q_p is the quantile value given by the quantile function $F^{-1}(q_p)$, based on the cdf above:

$$
\begin{aligned}
q_p &= 1 - exp(-\lambda (F^{-1}(q_p))^2) \\
exp(-\lambda (F^{-1}(q_p))^2) &= 1 - q_p \\
-\lambda (F^{-1}(q_p))^2 &= \ln(1 - q_p) \\
(F^{-1}(q_p))^2 &= \frac{-1}{\lambda}\ln(1 - q_p) \\
F^{-1}(q_p) &= \sqrt{\frac{-1}{\lambda}\ln(1 - q_p)}
\end{aligned}
$$

(d) Compute the mean, $E[Y]$, and variance, $Var(Y)$. **Hint:** Use integration by parts.

$$
\begin{aligned}
E[Y] &= \int_{-\infty}^{\infty} yp(y)dy \\
&= \int_{0}^{\infty} 2y^2\lambda exp(-\lambda y^2)dy = \int_{0}^{\infty} x^{\frac{1}{2}}\lambda exp(-\lambda x) \\
Var(Y) &= E\left[Y^2\right] - E[Y]\, E[Y]
\end{aligned}
$$

4. Given a realization $y_1, y_2, \ldots, y_n$ from your random variable Y in the previous problem, what is the maximum likelihood estimate for λ?

The likelihood function will be

$$
\begin{aligned}
\mathcal{L}(\lambda) &= p_{Y_1,\ldots,Y_n}(y_1, y_2, \ldots, y_n; \lambda) \\
&= \prod_{i=1}^{n} p_{Y_i}(y_i; \lambda), && \text{assuming } Y_i \text{ are independently distributed} \\
&= \prod_{i=1}^{n} \left(2y\lambda exp(-\lambda y_i^2)\right) \\
\ln(\mathcal{L}(\lambda)) = \ell(\lambda) &= \sum_{i=1}^{n} \ln\left(2y_i\lambda exp(-\lambda y_i^2)\right) \\
&= \sum_{i=1}^{n} \ln(2y_i) + \sum_{i=1}^{n} \ln(\lambda) - \sum_{i=1}^{n} \lambda y_i^2
\end{aligned}
$$

In order to find the maximum likelihood, we can find the maximum log likelihood since log is a monotonously increasing function. The first term above is a constant and we can find the maximum likelihood by setting the derivative with respect to λ of the log likelihood to zero.

$$
\begin{aligned}
\frac{d\ell(\lambda)}{d\lambda} &= \frac{d\sum_{i=1}^{n}\ln(2y_i)}{d\lambda} + \frac{d\sum_{i=1}^{n}\ln(\lambda)}{d\lambda} - \frac{d\sum_{i=1}^{n}\lambda y_i^2}{d\lambda} \\
&= 0 + n\frac{d\ \ln(\lambda)}{d\lambda} - \sum_{i=1}^{n} y_i^2 \sum_{i=1}^{n}\frac{d\lambda}{d\lambda} \\
&= n\frac{1}{\lambda} - \sum_{i=1}^{n} y_i^2 = 0 \\
\frac{n}{\lambda} &= \sum_{i=1}^{n} y_i^2 \\
\lambda &= \frac{n}{\sum_{i=1}^{n} y_i^2}, && \text{which is the maximum likely estimate for } \lambda.
\end{aligned}
$$

R Coding Part

4. Recall that if F is a cdf and $U \sim \text{Unif}(0,1)$, then $Y = F^{-1}(U)$ will be a random variable with cdf F. Write a function that uses this method to generate n random numbers from the distribution of Y in Problem 3 above (n and λ should be parameters to the function).
Hint: Look at the `runif` function in R for simulating uniform random varibles.

(a) Generate 10,000 realizations of the random variable Y with $\lambda = 2$.

(b) Plot a histogram of these numbers, using the `hist` function with option `freq = FALSE`.

(c) Use the `lines` command to plot the pdf of Y on top.

(d) Compute the sample mean and variance of your 10,000 realizations. Do they roughly match the values for $\mathrm{E}[Y]$ and $\mathrm{Var}(Y)$ you calculated above?

5. Now generate 20 realizations from Y with $\lambda = 2$. Plot the log likelihood function for this data. Plot a vertical line at the maximum likelihood estimate for lambda (using the equation you got in Problem 3).

References

[1] *Conditional Expectation*, www.statlect.com/fundamentals-of-probability/conditional-expectation.