W271-2 – Spring 2016 – **Lab 1**

Juanjo Carin, Kevin Davies, Ashley Levato, Minghu Song

January 14, 2016

Contents

Part I: Marginal, Joint, and Conditional Probabilities

Question 1

In a team of data scientists, 36 are expert in machine learning, 28 are expert in statistics, and 18 are awesome. 22 are expert in both machine learning and statistics, 12 are expert in machine learning and are awesome, 9 are expert in statistics and are awesome, and 48 are expert in either machine learning or statistics or are awesome. Suppose you are in a cocktail party with this group of data scientists and you have an equal probability of meeting any one of them.

1. What is the probability of meeting a data scientist who is an expert in both machine learning and statistics and is awesome?

Let N be the size of the team of data scientists (i.e., the sample size), and M, S, A the event that a data scientist is either a machine learning expert, a statistics expert, or awesome.

$$
\begin{aligned}
\Pr(\mathbf{M} \cap \mathbf{S} \cap \mathbf{A}) &= \Pr(M) + \Pr(S \cap A) - \Pr(M \cup (S \cap A)) \\
&= \Pr(M) + \Pr(S \cap A) - \Pr((M \cup S) \cap (M \cup A)) \\
&= \Pr(M) + \Pr(S \cap A) - ((\Pr(M \cup S) + \Pr(M \cup A) - \Pr((M \cup S) \cup (M \cup A)))) \\
&= \Pr(M) + \Pr(S \cap A) - (\Pr(M) - \Pr(S) + \Pr(M \cap S)) - (\Pr(M) - \Pr(A) + \Pr(M \cap A)) + \Pr(M \cup S \cup A) \\
&= \Pr(M \cap S) + \Pr(M \cap A) + \Pr(M \cap A) - \Pr(M) - \Pr(S) - \Pr(A) + \Pr(M \cup S \cup A) \\
&= \frac{22}{N} + \frac{12}{N} + \frac{9}{N} - \frac{36}{N} - \frac{28}{N} - \frac{18}{N} + \frac{48}{N} = \frac{(22 + 12 + 9) - (36 - 28 - 18) + 48}{N} = \frac{43 - 82 + 48}{N} \\
&= \frac{\mathbf{9}}{\mathbf{N}}
\end{aligned}
$$

2. Suppose you meet a data scientist who is an expert in machine learning. Given this information, what is the probability that s/he is not awesome?

$$
\Pr(\mathbf{A^c}|\mathbf{M}) = 1 - \Pr(A|M) = 1 - \frac{\Pr(A \cap M)}{\Pr(M)} = 1 - \frac{\frac{12}{N}}{\frac{36}{N}} = 1 - \frac{12}{36} = 1 - \frac{1}{3} = \frac{\mathbf{2}}{\mathbf{3}} = \mathbf{0.6667}
$$

3. Suppose the you meet a data scientist who is awesome. Given this information, what is the probability that s/he is an expert in either machine learning or statistics?

$$
\begin{aligned}
\Pr(\mathbf{M} \cup \mathbf{S}|\mathbf{A}) &= \frac{\Pr((M \cup S) \cap A)}{\Pr(A)} = \frac{\Pr((M \cap A) \cup (S \cap A))}{\Pr(A)} \\
&= \frac{\Pr(M \cap A) + \Pr(S \cap A) - \Pr((M \cap A) \cap (S \cap A))}{\Pr(S)} = \frac{\Pr(M \cap A) + \Pr(S \cap A) - \Pr(M \cap S \cap A)}{\Pr(S)} \\
&= \frac{\frac{12}{N} + \frac{9}{N} - \frac{9}{N}}{\frac{18}{N}} = \frac{12 + 9 - 9}{18} = \frac{12}{18} \\
&= \frac{\mathbf{2}}{\mathbf{3}} = \mathbf{0.6667}
\end{aligned}
$$

Question 2

Suppose for events A and B, $\Pr(A) = p \leq \frac{1}{2}, \Pr(B) = q$, where $\frac{1}{4} < q < \frac{1}{2}$. These are the only information we have about the events.

1. What are the maximum and minimum possible values for $\Pr(A \cup B)$?

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$$

The maximum value of $\Pr(A \cup B)$ occurs when $A \cap B$ is the smallest set possible. In this case, since $\Pr(A) \leq \frac{1}{2}$ and $\Pr(B) < \frac{1}{2}$, it would be $A \cap B = \varnothing$ (if A and B were disjoint sets, which might be the case), so:

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B) = \Pr(A) + \Pr(B) - \Pr(\varnothing) = p + q - 0 = p + q$$

which could have a maximum value close to 1 (i.e., $A \cup B \approx \Omega$), in case $\Pr(A) = \frac{1}{2}$ and $\Pr(B) \approx \frac{1}{2}$.

The minimum value of $\Pr(A \cup B)$ occurs when $A \cup B$ is the largest set possible, A or B. In this case, since $\Pr(A)$ does not have a lower bound, that would happen when $A \subseteq B$ and (consequently) $A \cap B = A$, which would lead to:

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A) = \Pr(B) = q$$

whose minimum value is greater than $\frac{1}{4}$.

In summary,

$$\mathbf{\frac{1}{4}} < \Pr(\mathbf{A} \cup \mathbf{B}) < \mathbf{1}$$

2. What are the maximum and minimum possible values for $\Pr(A|B)$?

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

If $B \subseteq A$ (which would imply that the lower bound for p is also $\frac{1}{4}$), then:

$$\Pr(A|B) = \frac{\Pr(B)}{\Pr(B)} = 1$$

As seen in the previous part, since $\Pr(A) \leq \frac{1}{2}$ and $Pr(B) < \frac{1}{2}$, it might occur that $A \cap B = \varnothing$, and hence:

$$\Pr(A|B) = \frac{\Pr(\varnothing)}{\Pr(B)} = \frac{0}{q} = 0$$

irrespective of the value of q.

In summary,

$$\mathbf{0} \leq \Pr(\mathbf{A}|\mathbf{B}) \leq \mathbf{1}$$

Part II: Random Variables, Expectation, Conditional Exp.

Question 3

Suppose the life span of a particular server is a continuous random variable, t, with a uniform probability distribution between 0 and k year, where $k \leq 10$ is a positive integer.

The server comes with a contract that guarantees a full or partial refund, depending on how long it lasts. Specifically, if the server fails in the first year, it gives a full refund denoted by θ. If it lasts more than 1 year but fails before $\frac{k}{2}$ years, the manufacturer will pay $x = \$A(k-t)^{1/2}$, where A is some positive constant equal to 2 if $t \leq \frac{k}{2}$. If it lasts between $\frac{k}{2}$ and $\frac{3k}{4}$ years, it pays $\frac{\theta}{10}$.

1. Given that the server lasts for $\frac{k}{4}$ years without failing, what is the probability that it will last another year?

2. Compute the expected payout from the contract, $E(x)$.

3. Compute the variance of the payout from the contract.

Question 4

Continuous random variables X and Y have a joint distribution with probability density function $f(x,y) = 2e^{-x}e^{-2y}$ for $0 < x < \infty, 0 < y < \infty$ and 0 otherwise.

1. Compute $\Pr(X > a, Y < b)$, where a, b are positive constants and $a < b$.

$$\Pr(X > a, Y < b) = \int_{x=a}^{\infty}\int_{y=0}^{b} f(x,y)dxdy$$

$$= \int_{x=a}^{\infty}\int_{y=0}^{b} 2e^{-x}e^{-2y}dxdy = 2\left(\int_{x=a}^{\infty} e^{-x}dx\right)\left(\int_{y=0}^{b} e^{-2y}dy\right)$$

$$= 2\left[-e^{-x}\right]_{x=a}^{\infty}\left[-\frac{e^{-2y}}{2}\right]_{y=0}^{b} = \left[e^{-x}\right]_{x=a}^{\infty}\left[e^{-2y}\right]_{y=0}^{b} = \left(0 - e^{-a}\right)\left(e^{-2b} - 1\right)$$

$$= \mathbf{e^{-a}\left(1 - e^{-2b}\right) = e^{-a} - e^{-a-2b}}$$

2. Compute $\Pr(X < Y)$.

$$\Pr(X < Y) = \int_{x=0}^{y}\int_{y=0}^{\infty} f(x,y)dxdy$$

$$= \int_{x=0}^{y}\int_{y=0}^{\infty} 2e^{-x}e^{-2y}dxdy = 2\int_{y=0}^{\infty}\left(\int_{x=0}^{y} e^{-x}dx\right)e^{-2y}dy$$

$$= 2\int_{y=o}^{\infty}\left[-e^{-x}\right]_{x=0}^{y} e^{-2y}dy = 2\int_{y=0}^{\infty}\left(1 - e^{-y}\right)e^{-2y}dy = 2\int_{y=0}^{\infty}\left(e^{-2y} - e^{-3y}\right)dy$$

$$= 2\left[-\frac{e^{-2y}}{2} + \frac{e^{-3y}}{3}\right]_{y=0}^{\infty} = 2\left[0 - \left(-\frac{1}{2} + \frac{1}{3}\right)\right] = 2\left(\frac{1}{2} - \frac{1}{3}\right) = 1 - \frac{2}{3}$$

$$= \mathbf{\frac{1}{3}}$$

3. Compute $\Pr(X < a)$.

$$\begin{aligned}
\Pr(X < a) &= \int_{x=0}^{a}\int_{y=0}^{\infty} f(x,y)dxdy \\
&= \int_{x=0}^{a}\int_{y=0}^{\infty} 2e^{-x}e^{-2y}dxdy = 2\left(\int_{x=0}^{a} e^{-x}dx\right)\left(\int_{y=0}^{\infty} e^{-2y}dy\right) \\
&= 2\left[-e^{-x}\right]_{x=0}^{a}\left[-\frac{e^{-2y}}{2}\right]_{y=0}^{\infty} = \left[e^{-x}\right]_{x=0}^{a}\left[e^{-2y}\right]_{y=0}^{\infty} = \left(e^{-a}-1\right)(0-1) \\
&= \mathbf{1-e^{-a}}
\end{aligned}$$

Question 5

Let X be a random variable and x be a real number. A linear function of the squared deviation from x is another random variable, $Y = a + b(X-x)^2$, where a and b are some positive constant.

1. Find the value of x that minimizes $E(Y)$. Show that your result is really the minimum.

The *Law of the unconscious statistician* states that:

$$E\left[g(X)\right] = \int g(x)f_X(x)dx$$

So, if we call $\mu = E(X)$ and $\sigma^2 = Var(X)$ (and knowing that $Var(X) = E\left(X^2\right) - \left(E(X)\right)^2 = E\left(X^2\right) - \mu^2$):

$$\begin{aligned}
E(Y) &= \int \left(a + b(X-x)^2\right) f(X)dX \\
&= \int \left(a + b(X^2 + x^2 - 2xX\right) f(X)dX \\
&= (a + bX^2)\int f(X)dX - 2bx\int Xf(X)dX + b\int X^2 f(X)dX \\
&= (a + bx^2)1 - 2bx\mu + b(\sigma^2 - \mu^2) \\
&= bx^2 - 2b\mu x + \left(a + b(\sigma^2 + \mu^2)\right)
\end{aligned}$$

Hence:

$$\frac{dE(Y)}{dx} = 2bx - 2b\mu = 2b(x-\mu) = 2b\left(x - E(X)\right)$$

And consequently:

$$\frac{dE(Y)}{dx} \Longrightarrow \mathbf{x = E(X)}$$

2. Find the value of $E(Y)$ for the choice of t you found in (1)?

N.B.: We assume this is typo and it really means x, not t.

We just have to substitute in the last expression of $E(Y)

$$\mathbf{E(Y)} = \int \left(a + b(X-\mu)^2\right) f(X)dX$$

$$= b\mu^2 - 2b\mu^2 + \big(a + b(\sigma^2 + \mu^2)\big) = -b\mu^2 + a + b\sigma^2 + b\mu^2$$
$$= \mathbf{a + b\sigma^2}$$

3. Suppose $Y = at + b(X - x)^2$. Find the values of x that minimizes $E(Y)$. Show that your result is really the minimum.

N.B.: Another typo? If $Y = at + b(X - x)^2$ the answer is the same than in (1). Let's use $Y = ax + b(X - x)^2$ instead.

$$E(Y) = \int \big(ax + b(X - x)^2\big)\, f(X)dX$$
$$= \int \big(ax + b(X^2 + x^2 - 2xX\big)\, f(X)dX$$
$$= (ax + bX^2) \int f(X)dX - 2bx \int Xf(X)dX + b \int X^2 f(X)dX$$
$$= (ax + bx^2)1 - 2bx\mu + b(\sigma^2 - \mu^2)$$
$$= bx^2 + (a - 2b\mu)x + b(\sigma^2 + \mu^2)$$

Hence:

$$\frac{dE(Y)}{dx} = 2bx + (a - 2b\mu) = 2b(x - \mu) + a = 2b\,(x - E(X)) + a$$

And consequently:

$$\frac{dE(Y)}{dx} \Longrightarrow \mathbf{x = E(X) - \frac{a}{2b}}$$

Question 6

Suppose X and Y are independent continuous random variables, where both of which are uniformly distributed between 0 and 1. Let random variable $Z = X + Y$.

1. Choose a value of z between 0 and 2, and draw a graph depicting the region of the $X - Y$ plane for which Z is less than z.

2. Derive the probability density function, $f(z)$.