# Lecture 7: Total Probability Law & Conditional Expectation

*23 March, 2018*

*Sunil Kumar Gauttam*

*Department of Mathematics, LNMIIT*

We state the law of total probability for continuous random variable, which is completely analogous to the discrete case.

**Theorem 7.1** *Let $X$ be a continuous random variable with the pdf $f_X$. Then for any event $B$,*

$$P(B) = \int_{-\infty}^{\infty} P(B|X = x) f_X(x)dx.$$

**Example 7.2** *Let $X$ and $Y$ be two independent uniform(0,1) random variables. Find $P(X^3 + Y > 1)$.*

**Solution:** We discuss two solution methods, which one makes our life simple, we shall see it.

1. One approach would be form the joint pdf $f(x, y) = f_X(x)f_Y(y)$ and then compute the following integral:
$$P(X^3 + Y > 1) = \iint_A f(x, y)dxdy,$$
where $A := \{(x, y) \in \mathbb{R}^2 | x^3 + y > 1\}$. Since $X$ and $Y$ are independent, the joint pdf of $X$ and $Y$ is $f(x, y) = f_X(x)f_Y(y)$, where
$$f_X(x) = \begin{cases} 1 & \text{if} \quad 0 < x < 1 \\ 0 & \quad\text{otherwsie} \end{cases}, \; f_Y(y) = \begin{cases} 1 & \text{if} \quad 0 < y < 1 \\ 0 & \quad\text{otherwsie} \end{cases},$$
Hence
$$f(x, y) = \begin{cases} 1 & \text{if} \quad 0 < x < 1, 0 < y < 1 \\ 0 & \quad\text{otherwsie} \end{cases}.$$
We need to plot the graph of the function $f(x) = 1 - x^3$. It's time to recall curve tracing from M-I.
    (a) $f$ being a polynomial is defined throughout $\mathbb{R}$. Since $f$ is a continuous and $\lim_{x\to\infty} f(x) = -\infty$, $\lim_{x\to-\infty} f(x) = \infty$, so by intermediate value property, range of $f$ is $(-\infty, \infty)$.

(b) $f'(x) = -3x^2 \leq 0$ for all $x \in \mathbb{R}$, therefore $f$ is a strictly decreasing function on $\mathbb{R}$.

(c) $f'(x) = 0$ tells us $x = 0$ is the only critical point. But as $f$ is strictly decreasing in neighborhood of 0, therefore $x = 0$ can neither be a point of local minima or local maxima.

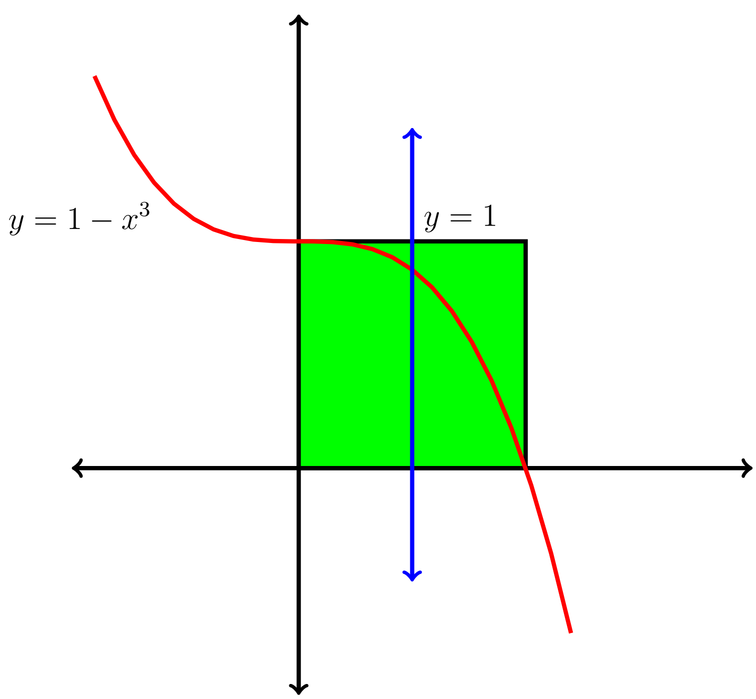
(d) $f''(x) = -6x$. So $f''(x) < 0$ in interval the $(0, \infty)$,i.e., $f$ is concave down in the interval $(0, \infty)$.

Also $f''(x) > 0$ in interval the $(-\infty, 0)$,i.e., $f$ is concave up (or convex) in the interval $(-\infty, 0)$.

Therefore $x = 0$ is the point of inflection.

(e) Graph of $f$ passes through point $(0, 1)$ and $(1, 0)$.

Gathering all the above information we plot the graph of $f(x) = 1 - x^3$.




Setting up the limits as in figure:

$$\iint\limits_A f(x, y)dxdy = \int_0^1 \left( \int_{1-x^3}^1 dy \right) dx = \int_0^1 \left(1 - (1 - x^3)\right) dx = \int_0^1 x^3 dx = \frac{1}{4}.$$

2. Now we illustrate how the conditioning is useful: One can condition either w.r.t. $Y = y$ or $X = x$.

(a) We condition w.r.t. $Y$. Hence by total probability law:

$$\begin{aligned}
P(X^3+Y>1) &= \int_{-\infty}^{\infty} P(X^3+Y>1|Y=y)f_Y(y)dy = \int_0^1 P(X^3+y>1|Y=y)dy \\
&= \int_0^1 P(X^3>1-y|Y=y)dy = \int_0^1 P(X^3>1-y)dy \\
&= \int_0^1 P(X>(1-y)^{\frac{1}{3}})dy \ (\because 0<y<1, X^3>1-y \iff X>(1-y)^{\frac{1}{3}}\ ) \\
&= \int_0^1 \left(\int_{\sqrt[3]{1-y}}^1 dx\right) dy \\
&= \int_0^1 \left[1-\sqrt[3]{1-y}\right] dy = \int_1^0 (1-\sqrt[3]{u})(-du) = \left[u-\frac{3}{4}u^{\frac{4}{3}}\right]_0^1 = \frac{1}{4}
\end{aligned}$$

(b) We condition w.r.t. $X$. Hence by total probability law:

$$\begin{aligned}
P(X^3+Y>1) &= \int_{-\infty}^{\infty} P(X^3+Y>1|X=x)f_X(x)dx = \int_0^1 P(X^3+Y>1|X=x)dx \\
&= \int_0^1 P(x^3+Y>1|X=x)dx = \int_0^1 P(Y>1-x^3|X=x)dx \\
&= \int_0^1 P(Y>1-x^3)dx = \int_0^1 \left(\int_{1-x^3}^1 dy\right) dx \\
&= \int_0^1 x^3 dx = \frac{x^4}{4}\Big|_0^1 = \frac{1}{4}
\end{aligned}$$

∎

# Conditional Expectation

Now we define conditional expectation denoted by $E[X|Y=y]$ of the random variable $X$ given the information $Y=y$. A conditional expectation is the same as an ordinary expectation, except that it refers to the new universe and pmf is replaced by their conditional counterparts.

**Definition 7.3** *Let $X$ and $Y$ be discrete random variables with conditional pmf $f_{X|Y}$ of $X$ given $Y$. Then conditional expectation of $X$ given $Y=y$ is defined as*

$$E[X|Y=y] = \sum_{x\in R_X} x f_{X|Y}(x|y)$$

**Example 7.4** *Let $X, Y$ be independent random variables with geometric distribution of parameter $0 < p < 1$. Calculate $E[Y|X+Y=n]$ where $n \geq 2$.*

**Solution:** Set $Z := X + Y$. Observe that

$$P(Y = y|Z = n) = P(\emptyset) = 0 \text{ if } y \geq n$$

For $y = 1, 2, \cdots, n-1$, we have computed earlier

$$P(Y = y|Z = n) = \frac{1}{n-1}$$

So

$$f_{Y|Z}(y|n) = \begin{cases} \frac{1}{n-1} & ; \quad y = 1, 2, \cdots, n-1 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

Hence

$$\begin{aligned} E[Y|Z = n] &= \sum_y y f_{Y|Z}(y|n) \\ &= \sum_{y=1}^{n-1} y \frac{1}{n-1} \\ &= \frac{1}{n-1} \times \frac{(n-1)(n-1+1)}{2} \\ &= \frac{n}{2} \end{aligned}$$

■

**Definition 7.5** *Let $X$ and $Y$ be random variables with conditional pdf $f_{X|Y}$ of $X$ given $Y$. The conditional expectation of $X$ given $\{Y = y\}$ is defined as*

$$E[X|Y = y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y) dx$$

**Theorem 7.6**

1. *Let $X, Y$ be discrete random variables with joint pmf $f$, marginal pmfs $f_X$ and $f_Y$ respectively. If $Y$ has finite mean then*

$$E[Y] = \sum_x E[Y|X = x] f_X(x)$$

2. *Let $X, Y$ be random variables with joint pdf $f$, marginal pdfs $f_X$ and $f_Y$ respectively. If $Y$ has finite mean then*

$$E[Y] = \int_{-\infty}^{\infty} E[Y|X = x] f_X(x) dx$$

**Proof:**

$$\begin{aligned}
\sum_x E[Y|X=x]f_X(x) &= \sum_x \left( \sum_y y f_{Y|X}(y|x) \right) f_X(x) \\
&= \sum_x \sum_y y f(x,y) \\
&= \sum_y y \sum_x f(x,y) \\
&= \sum_y y f_Y(y) \\
&= EY
\end{aligned}$$

■

**Remark 7.7** *The above theorem is called total expectation theorem. It express the fact that "the unconditional average can be obtained by averaging the conditional averages". They can be used to calculate the unconditional expectation $E[X]$ from the conditional expectation.*