Conditional Distributions and Independence

30M

Conditional probability is the probability that an event occurs, **given** that another event has occurred. In probability notation, the probability that event A occurs given that event B has occurred is:

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

The numerator can be converted to:

$$\Pr(A \cap B) = \Pr(B|A) \cdot \Pr(A)$$

Combining the first and second equations above produces *Bayes' Theorem*.

$$\Pr(A|B) = \frac{\Pr(B|A)\Pr(A)}{\Pr(B)} \tag{S2.1.5.1}$$

The concept of conditional probabilities can be extended to random variables to derive *conditional distributions*. If a random variable X is conditioned on values between j and k, then its conditional PDF is

$$f_{X\,|\,j<X<k}(x) = \frac{f_X(x)}{\Pr(j < X < k)}, \text{ where } j < x < k$$

Using the conditional PDF, the following can be calculated:

- $\Pr(X \le x \mid j < X < k) = \int_j^x f_{X\,|\,j<X<k}(t)\,\mathrm{d}t$

- $\mathrm{E}[X \mid j < X < k] = \int_j^k x \cdot f_{X \mid j<X<k}(x)\,\mathrm{d}x$

For a discrete random variable X, replace the integrals and the PDFs in the equations above with sums and PMFs, respectively.

A random variable can also be conditioned on **another** random variable. For example, if a continuous random variable X is conditioned on another random variable Y, then this new conditional random variable is denoted $(X \mid Y)$. Its conditional probability and conditional mean can be determined as follows:

- Conditional probability: $\Pr(X \le c \mid Y = y) = \int_{-\infty}^{c} f_{X \mid Y}(x \mid y)\,\mathrm{d}x$

- Conditional mean: $\mathrm{E}[X \mid Y = y] = \int_{-\infty}^{\infty} x \cdot f_{X \mid Y}(x \mid y)\,\mathrm{d}x$

Again, if $(X \mid Y)$ is discrete, replace the integrals and the PDFs in the equations above with sums and PMFs, respectively.

The concept of conditional distributions is very important in relation to the three theorems below.

The Law of Total Probability

The ***Law of Total Probability*** states that if events A_1, A_2, ..., and A_n are mutually exclusive and form a partition of a sample space, then

$$\begin{aligned}\Pr(B) &= \sum_{i=1}^{n} \Pr(B \cap A_i) \\ &= \sum_{i=1}^{n} \left[\Pr(B \vert A_i) \cdot \Pr(A_i)\right] \qquad \text{(S2.1.5.2)}\end{aligned}$$

The same concept can be applied to determine the unconditional PMF of a discrete random variable X from a conditional distribution. In general,

$$\Pr(X = x) = \mathrm{E}_Y[\Pr(X = x \mid Y)] \qquad \textbf{(S2.1.5.3)}$$

Loosely speaking, the random variable Y takes the place of A_i in (S2.1.5.2). The subscript Y indicates that Y is the random variable for the "averaging". In other words,

- if Y is **discrete**, then

$$\Pr(X = x) = \sum_{\text{all } y} [\Pr(X = x \mid Y = y) \cdot \Pr(Y = y)]$$

- if Y is **continuous**, then

$$\Pr(X = x) = \int_{-\infty}^{\infty} \Pr(X = x \mid Y = y) \cdot f_Y(y)\,\mathrm{d}y$$

Note that X is **discrete** in both cases above. While this formula can be generalized to determine the unconditional PDF of a continuous random variable, the discrete form is more commonly used.

In addition, the Law of Total Probability can be used to determine the unconditional CDF and survival function of a random variable.

$$F_X(x) = \mathrm{E}_Y[\Pr(X \le x \mid Y)]$$

$$S_X(x) = \mathrm{E}_Y[\Pr(X > x \mid Y)]$$

The Law of Total Expectation

The ***Law of Total Expectation*** states that:

$$\mathrm{E}_X[X] = \mathrm{E}_Y[\mathrm{E}_X[X \mid Y]] \qquad (\text{S2.1.5.4})$$

Keep in mind that the inner expectation, $\mathrm{E}_X[X \mid Y]$, is a function of Y, not a constant. Remember that $(X \mid Y)$ is in terms of both X and Y. If Y changes, the conditional distribution will change, and so will the conditional mean.

More generally, for any function $g(X)$,

$$\mathrm{E}_X[g(X)] = \mathrm{E}_Y[\mathrm{E}_X[g(X) \mid Y]]$$

The Law of Total Variance

The ***Law of Total Variance*** states that:

$$\mathrm{Var}_X[X] = \mathrm{E}_Y[\mathrm{Var}_X[X \mid Y]] + \mathrm{Var}_Y[\mathrm{E}_X[X \mid Y]] \qquad (\text{S2.1.5.5})$$

The proof of this formula is lengthy; it is included in the appendix at the end of this section.

Coach's Remarks

Many students assume that the unconditional variance equals the average of the conditional variances. This is incorrect because the variance of the conditional mean also needs to be included in the calculation.

$$\begin{aligned}\mathrm{Var}_X[X] &= \mathrm{E}_Y[\mathrm{Var}_X[X \mid Y]] + \mathrm{Var}_Y[\mathrm{E}_X[X \mid Y]] \\ &\neq \mathrm{E}_Y[\mathrm{Var}_X[X \mid Y]]\end{aligned}$$

Let's apply these theorems in the examples below.

Example S2.1.5.1

For a health insurance policy, the insureds are classified into two classes: smoker and non-smoker. Of the pool of insureds, 40% are smokers and 60% are non-smokers.

The means and variances of the annual claim size for the insureds are:

Class	Mean	Variance
Smoker	600	1,000
Non-smoker	450	800

Determine

1. the expected annual claim size of a randomly selected insured.
2. the variance of the annual claim size of a randomly selected insured.

Solution to (1)

Let X be the annual claim size.

The question asks for the unconditional mean and variance. The means and variances given in the table are conditional on class.

Class	$\mathbf{E}[\boldsymbol{X} \mid \mathbf{Class}]$	$\mathbf{Var}[\boldsymbol{X} \mid \mathbf{Class}]$	$\mathbf{Pr}(\mathbf{Class})$
Smoker	600	$1{,}000$	0.4
Non-smoker	450	800	0.6

Because we are calculating the unconditional mean and variance from conditional means and variances, $(\text{S}2.1.5.4)$ and $(\text{S}2.1.5.5)$ should be used.

To determine the expected annual claim size, apply the Law of Total Expectation, i.e., $(\text{S}2.1.5.4)$.

$$\begin{aligned}\mathrm{E}[X] &= \mathrm{E}[\mathrm{E}[X \mid \text{Class}]] \\ &= \sum_{\text{all classes}} \mathrm{E}[X \mid \text{Class}] \cdot \Pr(\text{Class}) \\ &= 600(0.4) + 450(0.6) \\ &= \mathbf{510}\end{aligned}$$

Solution to (2)

To determine the variance of the annual claim size, apply the Law of Total Variance, i.e., (S2.1.5.5).

$$\mathrm{Var}[X] = \mathrm{E}[\mathrm{Var}[X \mid \text{Class}]] + \mathrm{Var}[\mathrm{E}[X \mid \text{Class}]]$$

Determine each component individually.

- Calculate $\mathrm{E}[\mathrm{Var}[X \mid \text{Class}]]$.

$$\begin{aligned}\mathrm{E}[\mathrm{Var}[X \mid \text{Class}]] &= \sum_{\text{all classes}} \mathrm{Var}[X \mid \text{Class}] \cdot \Pr(\text{Class}) \\ &= 1{,}000(0.4) + 800(0.6) \\ &= 880\end{aligned}$$

- Calculate $\mathrm{Var}[\mathrm{E}[X \mid \text{Class}]]$.

$$\mathrm{Var}[\mathrm{E}[X \mid \text{Class}]] = \mathrm{E}\left[\mathrm{E}[X \mid \text{Class}]^2\right] - \mathrm{E}[\mathrm{E}[X \mid \text{Class}]]^2$$

$$\begin{aligned} \mathrm{E}\left[\mathrm{E}[X \mid \text{Class}]^2\right] &= \sum_{\text{all classes}} \mathrm{E}[X \mid \text{Class}]^2 \cdot \Pr(\text{Class}) \\ &= 600^2(0.4) + 450^2(0.6) \\ &= 265{,}500 \end{aligned}$$

$$\mathrm{E}[\mathrm{E}[X \mid \text{Class}]] = 510$$

$$\begin{aligned} \mathrm{Var}[\mathrm{E}[X \mid \text{Class}]] &= 265{,}500 - 510^2 \\ &= 5{,}400 \end{aligned}$$

Therefore,

$$\begin{aligned} \mathrm{Var}[X] &= 880 + 5{,}400 \\ &= \mathbf{6{,}280} \end{aligned}$$

Coach's Remarks

Instead of using the Law of Total Variance, $\mathrm{Var}[X]$ can also be calculated using $\mathrm{E}\left[X^2\right] - \mathrm{E}[X]^2$, where the first and second moments of X are computed using the Law of Total Expectation.

$$\mathrm{E}\left[X^k\right] = \mathrm{E}\left[\mathrm{E}\left[X^k \mid \text{Class}\right]\right]$$

Example S2.1.5.2

The claim size of a homeowner's policy is distributed as follows:

Claim Size	Probability
50	0.3
100	$0.7-q$
150	q

q varies by insured. For each insured, q has a distribution with the following density function:

$$f(q)=\frac{q}{0.175},\qquad 0.1<q<0.6$$

Calculate

1. the expected claim size of a randomly selected insured.
2. the probability that a randomly selected insured has a claim of 100.

Solution to (1)

Let X be the claim size.

While not stated explicitly, the table in the question actually describes the distribution of $(X\mid q)$, **not** X. This is because the probabilities of the claim size depend on the values of q, where q varies. This fact is important to ensure correct application of the formulas.

To determine the expected claim size, apply (S2.1.5.4).

$$\mathrm{E}_X[X]=\mathrm{E}_q[\mathrm{E}_X[X\mid q]]$$

$$\begin{aligned}\mathrm{E}_X[X \mid q] &= \sum_{\text{all } x} x \cdot \Pr(X = x \mid q) \\ &= 50(0.3) + 100(0.7 - q) + 150q \\ &= 85 + 50q\end{aligned}$$

$$\begin{aligned}\mathrm{E}_X[X] &= \mathrm{E}_q[85 + 50q] \\ &= 85 + 50\mathrm{E}_q[q] \\ &= 85 + 50\int_{0.1}^{0.6} q \cdot \frac{q}{0.175}\,\mathrm{d}q \\ &= 85 + 50\int_{0.1}^{0.6} \frac{q^2}{0.175}\,\mathrm{d}q \\ &= 85 + 50\left[\frac{q^3}{0.525}\right]_{0.1}^{0.6} \\ &= 85 + 50\left(\frac{43}{105}\right) \\ &= \mathbf{105.48}\end{aligned}$$

■

Solution to (2)

To determine the probability of getting a claim of 100, apply $(\mathrm{S}2.1.5.3)$.

$$\begin{aligned}\Pr(X = 100) &= \mathrm{E}_q[\Pr(X = 100 \mid q)] \\ &= \mathrm{E}_q[0.7 - q] \\ &= 0.7 - \mathrm{E}_q[q] \\ &= 0.7 - \frac{43}{105} \\ &= \mathbf{0.2905}\end{aligned}$$

Coach's Remarks

The three properties of expected values discussed in Section S2.1.4 were applied in the solution above. Alternatively, we may use (S2.1.4.1):

$$\mathrm{E}\left[g(X)\right] = \int_{-\infty}^{\infty} g(x) \cdot f_X(x)\,\mathrm{d}x$$

$$\mathrm{E}\left[85 + 50q\right] = \int_{0.1}^{0.6} (85 + 50q) \cdot \frac{q}{0.175}\,\mathrm{d}q$$

$$\mathrm{E}\left[0.7 - q\right] = \int_{0.1}^{0.6} (0.7 - q) \cdot \frac{q}{0.175}\,\mathrm{d}q$$

While the final answers will be the same, the calculations will be more tedious. Therefore, you are encouraged to apply the properties of expected values to simplify the calculations as much as possible.

Independence

If events A and B are ***independent***, then the occurrence of A will not affect the probability of the occurrence of B, and vice versa. In mathematical terms,

$$\Pr(A \mid B) = \Pr(A)$$

$$\Pr(B \mid A) = \Pr(B)$$

Recall that in general,

$$\Pr(A \cap B) = \Pr(A \mid B) \cdot \Pr(B)$$

Therefore, with independence,

$$\Pr(A \cap B) = \Pr(A) \cdot \Pr(B) \qquad \textbf{(S2.1.5.6)}$$

The same concept can also be applied to random variables. If X and Y are independent, then the following holds true.

- Discrete X and Y:

$$\Pr(X = x \mid Y = y) = \Pr(X = x)$$

$$\Pr(X = x,\, Y = y) = \Pr(X = x) \cdot \Pr(Y = y)$$

- Continuous X and Y:

$$f_{X|Y}(x \mid y) = f_X(x)$$

$$f_{X,Y}(x,\, y) = f_X(x) \cdot f_Y(y)$$

Furthermore, if X and Y are independent, then

$$\mathrm{E}[g(X) \cdot h(Y)] = \mathrm{E}[g(X)] \cdot \mathrm{E}[h(Y)] \qquad \textbf{(S2.1.5.7)}$$

This leads to

$$\begin{aligned}
\mathrm{Cov}\left[X,\, Y\right] &= \mathrm{E}\left[XY\right] - \mathrm{E}\left[X\right]\mathrm{E}\left[Y\right] \\
&= \mathrm{E}\left[X\right]\mathrm{E}\left[Y\right] - \mathrm{E}\left[X\right]\mathrm{E}\left[Y\right] \\
&= 0
\end{aligned}$$