

Class: FY BSc

Subject : Probability & Statistics - 1

Chapter: Unit 4 Chapter 2

Chapter Name: Conditional Expectation

Today's Agenda

1. Conditional Expectation
1. The Random variable $E[Y|X]$
 1. Mean
 2. Variance
2. Compound Distributions
 1. Moments of Compound Distribution
 2. Generating Functions of Compound Distribution

1 Conditional Expectation



The conditional expectation of Y given $X = x$ is the mean of the conditional distribution of Y given $X = x$. This mean is denoted $E[Y|X = x]$, or just $E[Y|x]$.

For a discrete distribution, this will be:

$$E[Y|X = x] = \sum_i y_i P[Y = y_i|X = x] = \sum_i y_i \frac{P[Y=y_i, X=x]}{P[X=x]}$$

For a continuous distribution,

$$E[Y|X = x] = \int_y y \cdot f(y|x) dy = \int_y y \cdot \frac{f(x,y)}{f(x)} dy$$

2.1 The Random Variable $E[Y|X]$

The conditional expectation $E[Y | X = x] = g(x)$, say, is, in general, a function of x . It can be thought of as the observed value of a random variable $g(X)$. The random variable $g(X)$ is denoted $E[Y | X]$.

$E[Y|X]$ is also referred to as the regression of Y on X .

$E[Y | X]$, like any other function of X , has its own distribution, whose properties depend on those of the distribution of X itself.



Theorem: $E[E[Y | X]] = E[Y]$

$$\begin{aligned}\text{Proof: } E[E[Y | X]] &= \int E[Y|x] f_X(x) dx \\ &= \int (\int y f(y|x) dy) f_X(x) dx \\ &= \iint y f(x, y) dx dy \\ &= E[Y]\end{aligned}$$

2.2 The random variable $\text{var}[Y | X]$ and the “ $E[V] + \text{var}[E]$ ” result

The variance of the conditional distribution of Y given $X = x$ is denoted $\text{var}[Y | x]$, where:

$$\text{var}[Y | x] = E[\{Y - E[Y | x]\}^2 | x] = E[Y^2 | x] - (E[Y | x])^2$$

$\text{var}[Y | x]$ is the observed value of a random variable $\text{var}[Y | X]$ where:

$$\text{var}[Y | X] = E[Y^2 | x] - (E[Y | x])^2 = E[Y^2 | X] - \{g(X)\}^2$$

Hence $E[\text{var}[Y | X]] = E[E[Y^2 | X] - E[\{g(X)\}^2]] = E[Y^2] - E[\{g(X)\}^2]$ and so:

$$E[Y^2] = E[\text{var}[Y | X]] + E[\{g(X)\}^2]$$

So the variance of Y , $\text{var}[Y] = E(Y^2) - [E(Y)]^2$, is given by:

$$E[\text{var}(Y | X)] + E[\{g(X)\}^2] - [E\{g(X)\}]^2 = E[\text{var}(Y | X)] + \text{var}[g(X)]$$

i.e. **$\text{var}[Y] = E[\text{var}[Y | X]] + \text{var}[E[Y | X]]$.**





Question

CT3 September 2015 Q7

X and Y are discrete random variables with joint distribution given below.

	$Y = -1$	$Y = 0$	$Y = 1$
$X = 1$	0	$1/4$	0
$X = 0$	$1/4$	$1/4$	$1/4$

- (i) Determine the conditional expectation $E[Y|X = 1]$.
- (ii) Determine the conditional expectation $E[X|Y = y]$ for each value of y .
- (iii) Determine the expected value of X based on your conditional expectation results from part (ii).

Solution

$$(i) \quad E[Y|X=1] = \frac{-1 \times 0 + 0 \times \frac{1}{4} + 1 \times 0}{\frac{1}{4}} = 0$$

$$(ii) \quad E[X|Y=-1] = \frac{1 \times 0 + 0 \times \frac{1}{4}}{\frac{1}{4}} = 0, \quad E[X|Y=0] = \frac{1 \times \frac{1}{4} + 0 \times \frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}$$

$$E[X|Y=1] = \frac{1 \times 0 + 0 \times \frac{1}{4}}{\frac{1}{4}} = 0$$

Solution

$$(iii) \quad E(X) = E[E[X|Y]] = E[X|Y=-1] \times P(Y=-1)$$

$$+ E[X|Y=0] \times P(Y=0)$$

$$+ E[X|Y=1] \times P(Y=1)$$

$$\text{and } E(X) = 0 \times \frac{1}{4} + \frac{1}{2} \times \frac{1}{2} + 0 \times \frac{1}{4} = \frac{1}{4}$$

3 Compound Distributions



Let $S = X_1, X_2, \dots, X_N$ (and $S = 0$ if $N = 0$) where the X_i 's are independent, identically distributed (as a variable X) and are also independent of N . S is said to have a compound distribution.

S is the sum of a random number of random quantities, which is the defining feature of a compound distribution.

Illustration: N is the number of claims which arise in a portfolio of business and X_i is the amount of the i 'th claim. S is the total claim amount.

3 Compound Distributions

Compound distributions arise commonly in general insurance examples, the random variable N is often referred to as the “number of claims” and the distribution of the random variables $X_1, X_2, \dots, X_N$ is referred to as the “individual claim size distribution”, even where the compound distribution arises in another context.

To define a compound distribution, you need to know two components:

- the distribution of N (which is a discrete distribution) and
- the distribution of the X_i 's (which may be any distribution).

When the particular distribution of N is known, a compound distribution is referred to by the name of this distribution eg a compound Poisson distribution.

3.1 Moments of Compound Distributions

The mean and variance of S are easily found.

$$E(S|N = n) = E(X_1 + X_2 + \dots + X_N | N = n) = E(X_1 + X_2 + \dots + X_n) = n.E(X)$$

Similarly:

$$\text{var}(S|N = n) = n.\text{var}(X)$$

Therefore we have:

$$\mathbf{E(S) = E[E(S|N)] = E[NE(X)] = E(N).E(X)}$$

and

$$\begin{aligned} \text{Var}(S) &= E[\text{var}(S|N)] + \text{var}[E(S|N)] \\ &= E[N \text{var}(X)] + \text{var}[N E(X)] \\ &= \mathbf{E(N).var(X) + var(N) [E(X)]^2} \end{aligned}$$



Question

CT3 April 2015 Q3

Assume that in a large portfolio of insurance contracts the claim size is a normally distributed random variable with expected value 1000. Also assume that the number of claims is a random variable following a Poisson distribution with parameter $\lambda = 400$.

- (i) Calculate the expected value of the total claim amount from contracts in this portfolio.
- (ii) Calculate a lower limit for the standard deviation of the total amount of claims from contracts in this portfolio.

Solution

Let X be the size of an individual claim, and N be the number of claims.

(i) Expected total amount is $E[X]E[N] = 1,000 \times 400 = 400,000$

(ii) $\text{Var}(\text{total amount}) = E[N]V[X] + V[N]E[X]^2$

A lower bound for the variance is then obtained by assuming $V[X] = 0$, that is,

$$\text{STD}(\text{total amount}) = \sqrt{V[N]E[X]^2} = 20 \times 1000 = 20,000$$

3.2 Generating Function of Compound Distributions

The MGF of S is given by:

$$M_S(t) = E(e^{tS}) = E[E(e^{tS} \mid N)]$$

$$\begin{aligned} E(e^{tS} \mid N) &= E[\exp\{t(X_1, X_2, \dots, X_N)\} \mid N = n] \\ &= E[\exp\{t(X_1, X_2, \dots, X_n)\}] \\ &= \prod E[\exp(tX_i)] \\ &= [M_X(t)]^n \end{aligned}$$

Therefore we have:

$$M_S(t) = E[\{M_X(t)\}^N] = E[\exp\{N \cdot \log M_X(t)\}] = M_N\{\log M_X(t)\}$$



Question

CT3 September 2010 Q8

A certain type of claim amount (in units of £1,000) is modelled as an exponential random variable with parameter $\lambda = 1.25$. An analyst is interested in S , the total of 10 such independent claim amounts. In particular he wishes to calculate the probability that S exceeds £10,000.

- (i) (a) Show, using moment generating functions, that:
- (1) S has a gamma distribution, and
 - (2) $2.5S$ has a χ^2_{20} distribution.
- (b) Use tables to calculate the required probability.

Solution

- (i) (a) (1) Let X_i be a claim amount.

$$\text{Mgf of } X_i \text{ is } M_X(t) = \left(1 - \frac{t}{1.25}\right)^{-1}$$

$$\text{Mgf of } S = \sum_{i=1}^{10} X_i \text{ is } M_S(t) = [M_X(t)]^{10} = \left(1 - \frac{t}{1.25}\right)^{-10},$$

which is the mgf of a gamma(10, 1.25) variable.

$$(2) \quad \text{Mgf of } 2.5S \text{ is } E[e^{t(2.5S)}] = E[e^{(2.5t)S}] = M_S(2.5t) = (1 - 2t)^{-10},$$

which is the mgf of a gamma(10, 1/2) variable, i.e. χ_{20}^2 .

$$(b) \quad P(\text{total} > \text{£}10,000) = P(S > 10) = P(\chi_{20}^2 > 25) = 1 - 0.7986 = 0.2014$$



Question

CT3 April 2010 Q9

- The number of claims, N , arising over a period of five years for a particular policy is assumed to follow a “Type 2” negative binomial distribution (as in the book of Formulae and Tables page 9)
- Each claim amount, X (in units of £1,000), is assumed to follow an exponential distribution with parameter λ independently of each other claim amount and of the number of claims.

Let S be the total of the claim amounts for the period of five years, in the case $k = 2$, $p = 0.8$ and $\lambda = 2$.

(i) Calculate the mean and the standard deviation of S based on the above assumptions.

Now assume that:

- N follows a Poisson distribution with parameter $\mu = 0.5$, that is, with the same mean as N above;
- X follows a gamma distribution with parameters $\alpha = 2$ and $\lambda = 4$, that is, with the same mean as X above.

(ii) Calculate the mean and the standard deviation of S based on these assumptions.

(iii) Compare the two sets of answers in (i) and (ii) above.

Solution

$$(i) \quad E[N] = \frac{k(1-p)}{p} = \frac{2(0.2)}{0.8} = 0.5 \quad \text{and} \quad V[N] = \frac{k(1-p)}{p^2} = \frac{2(0.2)}{0.8^2} = 0.625$$

$$E[X] = \frac{1}{\lambda} = \frac{1}{2} = 0.5 \quad \text{and} \quad V[X] = \frac{1}{\lambda^2} = \frac{1}{2^2} = 0.25$$

$$E[S] = E[N]E[X] = 0.5 \times 0.5 = 0.25, \text{ i.e. } \pounds 250$$

$$V[S] = E[N]V[X] + V[N]\{E[X]\}^2 = 0.5 \times 0.25 + 0.625 \times 0.5^2 = 0.28125$$

$$\therefore SD[S] = 0.530, \text{ i.e. } \pounds 530$$

Solution

$$(ii) \quad E[N] = V[N] = \mu = 0.5$$

$$E[X] = \frac{\alpha}{\lambda} = \frac{2}{4} = 0.5 \quad \text{and} \quad V[X] = \frac{\alpha}{\lambda^2} = \frac{2}{4^2} = 0.125$$

$$E[S] = E[N]E[X] = 0.5 \times 0.5 = 0.25, \text{ i.e. } \pounds 250$$

$$V[S] = E[N]V[X] + V[N]\{E[X]\}^2 = 0.5 \times 0.125 + 0.5 \times 0.5^2 = 0.1875$$

$$\therefore SD[S] = 0.433, \text{ i.e. } \pounds 433$$

Solution

- (iii) As expected the means are the same,
but the standard deviation in (i) is larger than that in (ii) due to the fact that both N and X have larger variances.