Lecture

INSTITUTE OF ACTUARIAL
& QUANTITATIVE STUDIES

Class: FY BSc

Subject : Probability & Statistics 1

Chapter: Unit 2 Chapter 1

Chapter Name: Theoretical Discrete Distributions

Today's Agenda

1. Introduction
    1. Definitions

2. The Discrete Distributions
    1. Uniform Distribution
    2. Bernoulli Distribution
    3. Binomial Distribution
    4. Geometric Distribution
    5. Negative Binomial Distribution
    6. Hypergeometric Distribution
    7. Poisson Distribution

1 Introduction

Calculating probability distributions takes time!

So far we've looked at how to calculate and use probability distributions, but wouldn't it be nice to have something **easier to work with**, or just **quicker to calculate**?

In this chapter, we'll show you some **special probability distributions** that follow very definite patterns.

Once you know these patterns, you'll be able to use them to **calculate probabilities, expectations, and variances in record time**.

We'll Introduce Discrete Probability Distributions.

1.1 Definitions

In statistics, when we talk about distributions we usually mean probability distributions

Definition (informal): A distribution is a function that shows the possible values for a variable and how often they occur.

Which Statistical Distributions you know? If any.

The distributions considered in this chapter are all models for the number of something – *eg* number of "successes", number of "trials", number of deaths, number of claims. The values assumed by the variables are integers from the set {0, 1, 2, 3, …} – such variables are often referred to as counting variables.

2.1 Uniform Distribution

The Probability function is p(x) = 1/k for x = 1,2,….,k and p(x) = 0 otherwise.

Probability measure: equal assignment (1/*k*) to all outcomes, *ie* all outcomes are equally likely.

Moments:

$$\mu = E[X] = \frac{(1+2+\cdots+k)}{k} = \frac{\frac{1}{2}k(k+1)}{k} = \frac{k+1}{2}$$

$$E[X^2] = \frac{(1^2+2^2+\cdots+k^2)}{k} = \frac{\frac{1}{6}k(k+1)(2k+1)}{k} = \frac{(k+1)(2k+1)}{6}$$

$$\Rightarrow \sigma^2 = \frac{k^2-1}{12}$$

2.2 Bernoulli Distribution

A Bernoulli trial is an experiment which has (or can be regarded as having) only two possible outcomes – *s* ("success") and *f* ("failure").

Sample space $S = \{s, f\}$. The words "success" and "failure" are merely labels – they do not necessarily carry with them the ordinary meanings of the words.

Probability measure: $P(\{s\}) = \theta, P(\{f\}) = 1 - \theta$

Random variable X defined by $X(s) = 1, X(f) = 0$. X is the number of successes that occur (0 or 1).

Distribution:

$$P(X = x) = \theta^x (1-\theta)^{1-x}, x = 0,1; 0 < \theta < 1$$

Moments:

$$\mu = \theta$$
$$\sigma^2 = \theta - \theta^2 = \theta(1-\theta)$$

2.3 Binomial Distribution

Consider a sequence of n Bernoulli trials as above such that:
(i) the trials are independent of one another, *ie* the outcome of any trial does not depend on the outcomes of any other trials

and:

(ii) the trials are identical, *ie* at each trial $P(\{s\}) = \theta$. Such a sequence is called a "sequence of n independent, identical, Bernoulli (θ) trials" or, for short, a "sequence of n Bernoulli (θ) trials".

2.3 Binomial Distribution

Sample space S: the joint set of outcomes of all n trials

Probability measure: as above (Bernoulli) for each trial

Random variable X is the number of successes that occur in the n trials.

Distribution: $P(X = x) = \binom{n}{x}\theta^x(1-\theta)^{n-x}, x = 0,1,2,\ldots,n; 0 < \theta < 1$

Moments:

$$\mu = n\theta$$
$$\sigma^2 = n\theta(1-\theta)$$

2.3 Binomial Distribution

A binomial experiment possesses the following properties:

1. The experiment consists of a fixed number, n, of identical trials.
2. Each trial results in one of two outcomes: success,S, or failure,
3. The probability of success on a single trial is equal to some value p (or ϑ) and remains the same from trial to trial. The probability of a failure is equal to $q = (1 - p)$ (or $(1 - \vartheta)$).
4. The trials are independent.
5. The random variable of interest is X, the number of successes observed during the n trials.

2.4 Geometric Distribution

Consider again a sequence of independent, identical Bernoulli trials with $P(\{s\}) = \theta$. The variable of interest now is **the number of trials that have to be performed until the first success occurs.** Because trials are performed one after the other and a success is awaited, this distribution is one of a class of distributions called ***waiting-time distributions.***

Random variable X: number of the trial on which the first success occurs.

Distribution: For $X = x$ there must be a run of $(x - 1)$ failures followed by a success,
so $P(X = x) = \theta(1-\theta)^{x-1}$, $x = 1,2,3,\ldots$ $(0 < \theta < 1)$

Moments:

$$\mu = \frac{1}{\theta}$$

$$\sigma^2 = \frac{(1-\theta)}{\theta^2}$$

2.4 Geometric Distribution

Another formulation of the geometric distribution is sometimes used. Let Y be the number of failures before the first success.

Then $P(Y = y) = \theta(1 - \theta)^y$, $y = 0,1,2,3, \ldots$ with mean $\mu = \frac{1-\theta}{\theta}$.

$Y = X - 1$, where X is defined as above.

Question

CT3 September 2018 Q3

A sports scientist is building a statistical model to describe the number of attempts a high jump athlete will have to make until she succeeds in clearing a certain height for the first time during an indoor sports event. For this model the scientist considers a geometric distribution with probability of success p. The cumulative distribution function of the geometric distribution is given as
$F_X(x) = 1 - (1-p)^x$, $x = 1, 2, 3, \ldots$

(i) (a) State the assumptions that the scientist needs to make for considering this distribution.
(b) Comment on the validity of the assumptions in part (i)(a). [3]

The athlete has tried n jumps without success.

(ii) (a) Determine the probability that the athlete will require more than x additional jumps to succeed in clearing the height.
(b) Comment on what the answer in part (ii)(a) means for the athlete. [3]
[Total 6]

Solution

(i)

(a) Needs to assume that each time the athlete tries she independently has the same probability p of passing the height, i.e. that attempts here are iid. [1]

(b) Given that the attempts are at the same event and on the same day, it is reasonable to assume that conditions are the same (independence) and that probability of success does not change. [2]

(ii) (a) If X is the corresponding random variable, we want:

$$P(X > x + n \mid X > n)$$ [0.5]

$$= \frac{P(X>x+n)}{P(X>n)} = \frac{(1-p)^{x+n}}{(1-p)^{n}} = (1-p)^{x} = P(X > x)$$ [1.5]

(b) The lack of success on the first n jumps is irrelevant – under this model the chances of success are not any better because there have been n attempts already. [1]

[Total 6]

2.5 Negative Binomial Distribution

This is a generalisation of the geometric distribution.
The random variable X is the number of the trial on which the k th success occurs, where k is a positive integer.

Distribution: $P(X = x) = \binom{x-1}{k-1}\theta^k(1-\theta)^{x-k} x = k, k+1, \dots; 0 < \theta < 1$

We say **that X has a Type 1 negative binomial (k, θ) distribution.**

This distribution satisfies the recurrence relationship:

$$P(X = x) = \frac{x-1}{x-k}(1-\theta)P(X = x-1)$$

Note that in applying this model, the value of k is known. In this model k must be a positive integer.

Moments:

$$\mu = \frac{k}{\theta}$$

$$\sigma^2 = \frac{k(1-\theta)}{\theta^2}$$

2.5 Negative Binomial Distribution

Type 2 negative binomial distribution

Another formulation of the negative binomial distribution is sometimes used.

Let Y be the number of failures before the k th success.

Then $P(Y=y) = \binom{k+y-1}{y} \theta^k (1-\theta)^y, y = 0,1,2,3,\ldots$, with mean $\mu = \frac{k(1-\theta)}{\theta}$. $Y = X - k$, where X is defined as above.

Question

CT3 April 2010 Question 9

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) with mean

$$E[N] = \frac{k(1-p)}{p} \text{ and variance } V[N] = \frac{k(1-p)}{p^2}.$$

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 $\lambda = 2$.

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

Solution

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

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

$E[S] \ = E[N]E[X] \ = 0.5 \times 0.5 \ = \ 0.25$, i.e. £250

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

$\therefore \ SD[S] = 0.530$, i.e. £530

2.6 Hypergeometric Distribution

This is the “finite population” equivalent of the binomial distribution, in the following sense.

Suppose objects are selected at random, one after another, without replacement, from a finite population consisting of k “successes” and N - k “failures”. The trials are not independent, since the result of one trial (the selection of a success or a failure) affects the make-up of the population from which the next selection is made.

The mean is given by $\mu = \frac{nk}{N}$, which parallels the " $\mu = n\theta$ " result for the binomial distribution - the initial proportion of successes here being $\frac{k}{N}$.

The binomial, with $\theta = \frac{k}{N}$, provides a good approximation to the hypergeometric in many situations.

2.6 Poisson Distribution

- This distribution models the number of events that occur in a specified interval of time, when the events occur one after another in time in a well-defined manner. This manner presumes that the events occur singly, at a constant rate, and that the numbers of events that occur in separate (ie non-overlapping) time intervals are independent of one another. These conditions can be described loosely by saying that the events occur "randomly, at a rate of … per …", and such events are said to occur according to a Poisson process.

- Another approach to the Poisson distribution uses arguments which appear at first sight to be unrelated to the above. Consider a sequence of binomial (n, Θ) distributions as $n \to \infty$ and $\Theta \to 0$ together, such that the mean $n\Theta$ is held constant at the value λ . The limit leads to the distribution of the Poisson variable, with parameter λ .

2.6 Poisson Distribution

Distribution: $P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}, x = 0,1,2,3, \dots ; \lambda > 0$

This distribution satisfies the recurrence relationship:

$$P(X = x) = \frac{\lambda}{x} P(X = x - 1)$$

If X has a Poisson distribution with parameter λ, then we can write $X \sim \mathrm{Poi}(\lambda)$.

2.6 Poisson Distribution

Moments:

Since the binomial mean is held constant at λ through the limiting process, it is reasonable to suggest that the distribution of X (the limiting distribution) also has mean λ. This is in fact the case.

The binomial variance is:

$$n\theta(1-\theta) = n\left(\frac{\lambda}{n}\right)\left(1-\frac{\lambda}{n}\right) = \lambda\left(1-\frac{\lambda}{n}\right) \to \lambda \text{ as } n \to \infty$$

This suggests that X has variance λ. This is in fact also the case.

$$\text{So } \mu = \sigma^2 = \lambda$$

2.6 Poisson Distribution

Good Approximation to the Binomial

- The Poisson distribution provides a very good approximation to the binomial when n is large and □ is small – typical applications have n = 100 or more and Θ = 0.05 or less.

- The approximation depends only on the product $n\ \Theta$ (= λ) – the individual values of n and Θ are irrelevant.

- So, for example, the value of $P(X = x)$ in the case n = 200 and Θ = 0.02 is effectively the same as the value of $P(X = x)$ in the case n = 400 and Θ = 0.01.

- When dealing with large numbers of opportunities for the occurrence of "rare" events (under "binomial assumptions"), the distribution of the number that occur depends only on the expected number.

2.6 Poisson Process

- When events are described as occurring “as a Poisson process with rate λ ” or “randomly, at a rate of λ per unit time” then the number of events that occur in a time period of length t has a Poisson distribution with mean λt .

Question

Subject C1, April 1994, Q7

For a certain type of insurance business, the number of claims per policy in a year has a Poisson distribution with mean 0.4

Consider a policy, which you know, has given rise to at least one claim in the last year. The probability that this policy has in fact given rise to exactly two claims in the least year is:

A 0.054
B 0.163
C 0.330
D 0.992

Solution

Answer B.

$$P(X \geq 1) = 1 - P(X = 0) = 1 - 0.6703 = 0.3297$$
$$P(X = 2) = 0.0536$$

Using $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}$, we get:

$$P(X = 2 \mid X \geq 1) = \frac{P(X \geq 1 \cap X = 2)}{P(X \geq 1)} = \frac{P(X = 2)}{P(X \geq 1)} = \frac{0.0536}{0.3297} = 0.163$$