MAT2377

Catalin Rada

Version May 12, 2009

Lectures

Summary of Lectures 2-3-4

- Computation of classical probabilities

- Rules and axioms, e.g. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

- Independence $P(A \cap B) = P(A)P(B)$

- Computation of conditional probabilities - see TB example

- Bayes' Theorem

Discrete Random Variables

Random Variable is **a number that goes with each outcome of an experiment**.

A *discrete* random variable is just a random variable with finite or countably infinite range. A *continuous* random variable is a random variable with an interval (either infinite or finite) of real numbers for its range! Example of DISCRETE: number of bits received in error; Examples of CONTINUOUS: time, temperature, weight.

If we have a probability rule also defined on the experiment, the prob. that the rv takes each of its values is called its **probability distribution.**

Notation for random variables (rv's)

- Capital Roman letters e.g. X, Y used to denote rv's.

- Corresponding lower case letters e.g. x, y used to denote *generic values taken* by rv.

- A rv is a way to define events: if X takes values $0, 1, 2, \ldots$ then we can define events $\{X = 0\}$, $\{X = 1\}$, $\{X = 2\}$,. . . etc.

- The probability distribution (mass) function is

$$f(x) = P\left(\{X = x\}\right) = P(X = x) = \ldots$$

Example 10:

1. Flip a fair coin. Assign $X = 1$ if Head occurs, $X = 0$ if Tail occurs. Then $P(X = 1) = \frac{1}{2}$, $P(X = 0) = \frac{1}{2}$ is the probability distribution function.

2. Roll 1 fair die. X is the number. Then $P(X = 1) = P(X = 2) = \cdots = P(X = 6) = \frac{1}{6}$ and $P(3 \leq X \leq 5) = P(X = 3) + P(X = 4) + P(X = 5) = \frac{1}{2}$.

3. Number of calls, X, is a random variable with possible values $0, 1, 2, \ldots$.

PROPERTIES of the probability distribution (mass) function f of a discrete rv X with possible range $\{x_1, \ldots, x_n\}$:

a) $f(x_i) \geq 0$ or all i;   b) $\sum_{i=1}^{n} f(x_i) = 1$;

c) $f(x_i) = P(X = x_i)$ for all i.

Let us discuss 3-18/page 78 from textbook.... i) it is evident that $f(x) \geq 0$ for all x; ii) $\sum_{x \geq 0} \frac{3}{4}(\frac{1}{4})^x = \frac{3}{4}\sum_{x \geq 0}(\frac{1}{4})^x = \frac{3}{4} \times \frac{1}{1-\frac{1}{4}} = 1$; a) $P(X = 2) = f(2) = \frac{3}{4} \times \frac{1}{4^2} = \frac{3}{64}$; b) $P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = f(0) + f(1) + f(2) = \frac{3}{4} + \frac{3}{4} \times \frac{1}{4} + \frac{3}{4} \times \frac{1}{4^2} = \ldots$; c) , d)....

Expected value

We define the mean of a discrete random variable X as

$$\mu = \mathrm{E}(X) = \sum_x x P(X = x),$$

where the sum extends over all values x the rv X takes.

The expectation is just a property of a probability distribution, but we can *interpret* it as a "long-run average".

Examples

- Expected roll of fair die: $P(Z=1) = P(Z=2) = \cdots = P(Z=6) = \frac{1}{6}$,

$$\mathrm{E}(Z) = \frac{1}{6} \times 1 + \frac{1}{6} \times 2 + \cdots = \frac{1}{6}(1+2+3+4+5+6) = \frac{21}{6} = 3.5 \,.$$

- In a gambling game a player can win \$3 with probability $\frac{1}{3}$ and lose \$1 with probability $\frac{2}{3}$ for each \$1 bet. Find expected value of the game. Hint: value of the game = win - bet.

  Solution: Expected value of the game $= 2 \times \frac{1}{3} + (-2) \times \frac{2}{3} = -\frac{2}{3}$

Three statisticians went out hunting, and came across a large deer. The first statistician fired, but missed, by a meter to the left. The second statistician fired, but also missed, by a meter to the right. The third statistician didn't fire, but shouted in triumph, "On the average we got it!"

Mean of a function of a rv

If $Y = h(X)$,

$$\mathrm{E}(Y) = \sum_x h(x)P(X = x) = \sum_x h(x)f(x) \,.$$

Examples: If Z is the number showing on a roll of a "fair" 6-sided die, find $\mathrm{E}(Z^2)$ and $\mathrm{E}\left([Z - 3.5]^2\right)$. Here $h(z) = z^2$ and $h(z) = (z - 3.5)^2$;

$$
\begin{aligned}
\mathrm{E}\left(Z^2\right) &= \sum_z z^2 P(Z = z) = 1^2 \times \frac{1}{6} + 2^2 \times \frac{1}{6} + \cdots \\
&= \frac{1}{6}(1^2 + 2^2 + \cdots + 6^2) = \frac{91}{6} = 15\frac{1}{6} \,.
\end{aligned}
$$

$$\mathrm{E}\left([Z - 3.5]^2\right) = \cdots = \frac{1}{6}\left[(-2.5)^2 + (-1.5)^2 + (-0.5)^2 + 0.5^2 + 1.5^2 + 2.5^2\right]$$

$$= \frac{6.25 + 2.25 + 0.25 + 0.25 + 2.25 + 6.25}{6} = \frac{17.5}{6} = 2\frac{5}{6}.$$

Recall that for this Z, $\mathrm{E}(Z) = 3.5$.

Hence the second example above could be written as $\mathrm{E}\left([Z - \mathrm{E}(Z)]^2\right)$.

This is in fact an important general quantity called the **variance** of Z.

Variance

Variance is defined as the *expected squared difference from the expectation*:

$$\sigma^2 = V(X) = \mathrm{Var}(X) \;\; = \;\; \mathrm{E}\left([X - \mathrm{E}(X)]^2\right) = \mathrm{E}(X^2) - (\mathrm{E}[X])^2$$

We also introduce the **standard deviation** of a rv X as

$$\sigma = SD(X) = \sqrt{\mathrm{Var}(X)}\,.$$

Variance and SD allow us to compare probability distributions: those with higher variance/SD are *more spread out about the expectation*.

- $\mathrm{E}(aX) = a\,\mathrm{E}(X);\ \mathrm{E}(X + a) = \mathrm{E}(X) + a$

- $\mathrm{Var}(aX) = a^2\,\mathrm{Var}(X);\ \mathrm{Var}(X + a) = \mathrm{Var}(X)$

Example: Let X and Y be a random variables with the following p.d.'s:

x	$P(X = x)$
-2	$1/5$
-1	$1/5$
0	$1/5$
1	$1/5$
2	$1/5$

y	$P(Y = y)$
-4	$1/5$
-2	$1/5$
0	$1/5$
2	$1/5$
4	$1/5$

Calculate expected values. Compare variances.

Solution: We have $\mathrm{E}(X) = \mathrm{E}(Y) = 0$ and $\mathrm{Var}(X) < \mathrm{Var}(Y)$,

Binomial Coefficient

$$\binom{n}{r} = \frac{n!}{(n-r)!r!}.$$

- $2! \times 4! = (1 \times 2) \times (1 \times 2 \times 3 \times 4) = 48$, but $(2 \times 4)! = 8! = 40320$.

- $\binom{5}{1} = \frac{5!}{1! \times 4!} = \frac{1 \times 2 \times 3 \times 4 \times 5}{1 \times (1 \times 2 \times 3 \times 4)} = \frac{5}{1} = 5$; In general: $\binom{n}{1} = n$. Also $\binom{n}{0} = 1$

- $\binom{6}{2} = \frac{6!}{2! \times 4!} = \frac{4! \times 5 \times 6}{2! \times 4!} = \frac{5 \times 6}{2} = 15$

- $\binom{27}{22} = \frac{27!}{22! \times 5!} = \frac{22! \times 23 \times 24 \times 25 \times 26 \times 27}{5! \times 22!} = \frac{23 \times 24 \times 25 \times 26 \times 27}{120}$

Binomial experiment

- A Bernoulli trial is a random experiment with two possible outcomes, "success" and "failure". Let p denote the probability of a success.

- A binomial experiment consists of n repeated independent Bernoulli trials, each with the same probability of success, p.

Examples:

- female/male births; satisfactory/defective items on prod'n line; sampling with replacement with two types of item.

Binomial distribution

If we have a binomial experiment (i.e. n independent events, each with prob p) then the number X of them that occur is a random variable with a binomial(n, p) distribution:

$$f(x) = P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}, \quad \text{for } x = 0, 1, 2, \ldots, n.$$

This often abbrev'd to "$X \sim \mathcal{B}(n, p)$". (X is binomial rv or has binomial distribution with parameters n and p)

- If $X \sim B(1, p)$ then $P(X = 0) = 1 - p$, $P(X = 1) = p$ and so

$$\mathrm{E}(X) = (1 - p) \times 0 + p \times 1 = p.$$

Expectation and Variance for Binomial

Recall: if $X \sim \mathcal{B}(n,p)$ then $P(X = x) = \binom{n}{x}p^x(1-p)^{n-x}$ for $x = 0, 1, 2, \ldots, n$. It can be shown that

$$\mathrm{E}(X) = \sum_{x=0}^{n} x \binom{n}{x} p^x(1-p)^{n-x} = \cdots = \boldsymbol{np}$$

and so

$$\mathrm{Var}(X) = \mathrm{E}\left[(X - np)^2\right] = \sum_{x=0}^{n}(x - np)^2 \binom{n}{x} p^x(1-p)^{n-x} = \boldsymbol{np(1-p)}\,.$$

Later we will see an easier way to derive these by interpreting X as a sum of other rv's.

Example 9:

Suppose that each sample of water has 10% of being polluted. If 12 samples are selected independently, then it is reasonable to model the number X of polluted samples as $\mathcal{B}(12, 0.1)$. Find

(a) $\mathrm{E}(X)$ and $\mathrm{Var}(X)$.

(b) $P(X = 3)$.

(c) $P(X \leq 3)$ (tables can be used).

Solution:

(a) If $X \sim \mathcal{B}(n, p)$ then $\mathrm{E}(X) = np$ and $\mathrm{Var}(X) = np(1 - p)$ so

$$\mathrm{E}(X) = 12 \times 0.1 = 1.2 \quad ; \quad \mathrm{Var}(X) = 12 \times 0.1 \times 0.9 = 1.08 \, .$$

(b) $P(X = 3) = \binom{12}{3}(0.1)^3(0.9)^9 \approx 0.0852.$

(c) $P(X \leq 3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) = \cdots$.
However, $P(X \leq 3)$ for $X \sim B(12, 0.1)$ is tabulated on p709(710!!) of text: ≈ 0.9744.

- Also, since $P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) \approx 0.8891$ from tables we see that $P(X = 3) = P(X \leq 3) - P(X \leq 2) \approx 0.9744 - 0.8891 = 0.0853$ (note rounding error).

Example:

- An airline sells 101 tickets for a flight with 100 seats. Each ticketed passenger is known to have a 0.97 pb. of showing up for flight. What is probability of 101 passengers showing up (and airline caught overbooking)? Make appropriate assumptions.

- Do the same problem if the airline sells 125 tickets.

Sol: Consider the 101 tickets as 101 Bernoulli (success/failure) trials with success meaning that the passenger holding the ticket shows up (so that p = 0.97). We need then to compute $P(X = 101)$, and we do it as follows: $P(X = 101) = \binom{101}{101}(0.97)^{101}(0.03)^0 = etc$

Geometric random variable

Consider a sequence of Bernoulli trials, with probability p of success at each step. The geometric random variable X denotes the number of steps before the first success occurs. The probability distribution is given by

$$f(x) = P(X = x) = (1 - p)^{x-1}p, \quad x = 1, \ldots$$

We will write $X \sim Geo(p)$. For this rv we have $E(X) = \frac{1}{p}$ and $V(X) = \frac{1-p}{p^2}$.

Example: A fair die is thrown until it shows six. What is the probability that 5 throws are required?

Sol: If 5 throws are required, we have to compute $P(X = 5)$, where X is geometric $Geo(1/6)$, so $= (1 - \frac{1}{6})^{5-1} \times \frac{1}{6} = etc.$

Poisson random variable

Let the number of "changes" that occur in a continuous interval (of time or space) be counted. We have a Poisson process of rate λ if:

(a) The number of changes occurring in nonoverlapping intervals are independent.

(b) The probability of exactly one change in a short interval of length h is approximately $\lambda \times h$.

(c) The probability of two or more changes in a sufficiently short interval is essentially 0.

Assume that an experiment satisfies the above properties. Let X be the number of changes in an interval of length 1. How to calculate p.m.f of X?

We partition the unit interval into n disjoint subintervals of length $1/n$. By condition (b) the probability of one change occurring in one small subinterval is approximately $\lambda \times 1/n$. By condition (c) the probability of two or more changes is essentially equal to 0. By condition (a) we have a sequence of n Bernoulli trials with probability $p = \lambda \times 1/n$. Therefore

$$f(x) = P(X = x) \approx \frac{n!}{x!(n-x)!} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x}$$

Letting $n \to \infty$ we get

$$P(X = x) = \exp(-\lambda)\frac{\lambda^x}{x!}$$

Let $X \sim$ Poisson, λ:

- The p.m.f. of X is

$$f(x) = P(X = x) = e^{-\lambda}\frac{\lambda^x}{x!}, \ \ x = 0, \ldots, \infty$$

- $\mathrm{E}(X) = \mu = \lambda$

- $\mathrm{Var}(X) = \sigma^2 = \lambda$

Example:

A traffic flow is typically modeled by a Poisson distribution. It is known that the traffic flowing through an intersection is 6 cars/minute on average.

1. What is the probability of no cars through the intersection within 30 seconds?

$Answer:$ 6 cars/min=3 cars/30 sec.... Thus $\lambda = 3$. Then compute $P(X = 0)$, and we get: $P(X = 0) = e^{-3}(\frac{3^0}{0!}) = e^{-3}(\frac{1}{1}) = e^{-3} = etc$

Example

Here's an example where the Poisson distribution was used in a maternity hospital to work out how many births would be expected during the night. The hospital had 3000 deliveries each year, so if these happened randomly around the clock 1000 deliveries would be expected between the hours of midnight and 8.00 a.m. This is the time when many staff are off duty and it is important to ensure that there will be enough people to cope with the workload on any particular night. The average number of deliveries per night is 1000/365, which is 2.74. From this average rate the probability of delivering 0,1,2, etc babies each night can be calculated using the Poisson distribution. Some probabilities are:

$$
\begin{aligned}
f(0) &\quad 2.74^0 \exp(-2.74) \,/\, 0! &&= 0.065 \\
f(1) &\quad 2.74^1 \exp(-2.74) \,/\, 1! &&= 0.177 \\
f(2) &\quad 2.74^2 \exp(-2.74) \,/\, 2! &&= 0.242 \\
f(3) &\quad 2.74^3 \exp(-2.74) \,/\, 3! &&= 0.221
\end{aligned}
$$

- On how many days in the year would 5 or more deliveries be expected? (Ans. 52)

- Over the course of one year, what is the greatest number of deliveries expected in any night? (Ans. 8)

Answer:

- If X is the number of babies during a night, then X has Poisson distribution with $\lambda = 2.74$. To compute $365 \times P(X \geq 5)$. WHY?

Exercise 1. Let us say that flaws in a black brand of tape occur on the average of one flow per 1200 meters. Assume a Poisson distribution. What is the p.d.f (or p.m.f.) of the rv X, where X is the number of flaws in a 4800-meter roll of black tape?

Sol: For 1200 meters we expect 1 flaw, FOR 4800 meters we DO expect $\lambda = \frac{4800 \times 1}{1200} = 4$. Hence the p.d.f. of our X is given by $f_X(x) = e^{-4} \times \frac{4^x}{x!}$, where $x \geq 0$. What is $P(X = 45)$? Is the probability of getting 45 flaws in a 4800-meter roll! We can compute then $P(X = 45) = e^{-4} \times \frac{4^{45}}{45!}$.

Exercise 2. A producer of chocolate Easter eggs has information that 2% of its eggs are salty! Assume a Poisson distribution. What is the chance that a truck of 1000 of these eggs contains at most 3 salty eggs?

Sol: For 100 eggs we expect 2 salty eggs; then for 1000 eggs we expect $\lambda = \frac{2 \times 1000}{100} = 20$. We have to compute $P(X \leq 3)$. We get:
$$P(X \leq 3) = e^{-20}\{\frac{20^0}{0!} + \frac{20^1}{1!} + \frac{20^2}{2!} + \frac{20^3}{3!}\} = e^{-20}\{1 + 20 + 200 + \frac{4000}{3}\} = etc$$

Exercise 3. At blackboard!!!!

ONE MORE PROBLEM FOR YOU FROM 2.7 (BAYES's Theorem)-see the textbook! To detect fraud in phone cards, there is a software that tracks the number of metropolitan areas where calls originate every day! The software found the following: 1% of the legitimate users originate calls from 2 or more metropolitan areas in a single day; 30% of fraudulent users originate calls from 2 or more metropolitan areas in a single day; the proportion of fraudulent users is 0.01%; Question: If a user originates calls from from 2 or more metropolitan areas in a single day, what is the chance (probability) that the user is fraudulent?

Sol: Set A be the event "a fraudulent user", and set B be the event "a user calls from 2 or more metropolitan areas in a single day". We need to find: $P(A|B)$. So: $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A')P(A')} = \frac{0.30 \times 0.0001}{0.30 \times 0.0001 + 0.01 \times 0.9999} = etc$