GM – Probabilités et Statistique

http://moodle.epfl.ch/course/view.php?id=18431

Lecture 4

- Review : random variables
- Expectation and variance
- Review : binomial distribution
- Poisson distribution
- Approximation of the binomial distribution by the Poisson distribution

Review : discrete RVs

- **Random Variable (RV)** : a real function defined on the sample space
  - RV : CAPITAL LETTERS ; specific values : lower case letters

- **Discrete RV** :
  1. probability mass function (pmf) : $p(x) = P(X = x)$
  2. cumulative distribution function (cdf) : $F(x) = P(X \leq x) = \sum_{i \leq x} p(i)$

- **Solving problems with RVs**
  1. *Identify the RV*
  2. Determine the *distribution* (loi/pmf) of the RV
  3. *Translate* the question
  4. *Respond* to the question

Expected value

- For a discrete RV X with pmf $p(x)$, we define the **expectation** (or the **mean**) by :

$$E[X] = \sum_{\substack{\text{all possible} \\ \text{values } x}} x p(x)$$

- Thus, it's a *weighted mean* of possible values of X, where the weights are $P(X = x)$
- It is also possible to calculate the expectation of a *function* of the (discrete) RV X (let's say $g(X)$) in the same fashion
- $g(X)$ is also a discrete RV, so to calculate $E[g(X)]$ it is sufficient to find its pmf (distribution) $p(g(x))$
- We should be able to deduce the distribution from that of X

Example

Example 4.1 Let X = sum of the numbers when we toss 2 fair dice (independently).

$E[X] =$

x	2	3	4	5	6	7	8	9	10	11	12	
$p(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	
$F(x)$	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{6}{36}$	$\frac{10}{36}$	$\frac{15}{36}$	$\frac{21}{36}$	$\frac{26}{36}$	$\frac{30}{36}$	$\frac{33}{36}$	$\frac{35}{36}$	$\frac{36}{36}$	$[= 1]$

[Due to the symmetry of the pmf of X, this is what we would have guessed without a probability course ! !]

Example

Example 4.2 Calculate $E[X^2]$ for the following pmf for X :

$$P(X = -1) = 0.2 \qquad P(X = 0) = 0.5 \qquad P(X = 1) = 0.3.$$

Solution : We define a new RV $Y = X^2$. Now find the distribution of Y directly :

$$\begin{aligned} P(Y = 1) &= P(X = -1) + P(X = 1) = 0.2 + 0.3 = 0.5 \\ P(Y = 0) &= P(X = 0) = 0.5. \end{aligned}$$

Thus $E[X^2] = E[Y] = 1(0.5) + 0(0.5) = \underline{\underline{0.5}}$

Note : $(E[X])^2 \boxed{\neq} E[X^2]$

$E[g(X)]$ again

- *Another way* to look at $E[g(X)]$:
  in noting that $g(X) = g(x)$ when $X = x$, it is reasonable to thing that $E[g(X)]$ should be a *weighted average* of the values of $g(x)$, with weights $P(X = x)$
- **Theorem :** For a discrete RV X taking on values x_i, $i \geq 1$, with respective probabilities $p(x_i)$, then for any real-valued function g we have

$$E[g(X)] = \sum_i g(x_i)p(x_i)$$

- For any pair of constants (a, b), $E[aX + b] = aE[X] + b$

Summaries of a distribution

- For a given RV X and its cumulative distribution function F (or pmf $p(x)$), it would be useful to *summarize the properties* of F in two or three measures
- One such measure is given by $E[X]$, the expected value of X, which tells us something about the 'central' value of the distribution
- However, it tells us nothing about the *variability* of X around the expectation

Example

- Consider the RVs W, Y, and Z :

$$
\begin{aligned}
W &= 0 \\
Y &= \begin{cases} -1 \text{ with probability } \frac{1}{2} \\ +1 \text{ with probability } \frac{1}{2} \end{cases} \\
Z &= \begin{cases} -100 \text{ with probability } \frac{1}{2} \\ +100 \text{ with probability } \frac{1}{2} \end{cases}
\end{aligned}
$$

- All 3 of these have the same expected value ($=$ *??*), *but the deviations between the different values* of Y are bigger than those of W, and smaller than those of Z

Variance and standard deviation

- Since we expect X to take on values around its mean $E[X]$, one way to measure variation would be to look at the average (absolute) distance between X and its mean $E[X]$, $E[| X - \mu |]$, where $\mu = E[X]$
- It is more mathematically convenient to consider instead the average *squared* distance between X and its mean
- For a RV X with mean μ, the **variance** of X is defined as :

$$Var(X) = E[(X - \mu)^2]$$

- We can establish an *alternative formula* for calculating $Var(X)$ (easier to use in practice) :

$$Var(X) = E[X^2] - (E[X])^2$$

- The **standard deviation** of X (σ) is the square root of $Var(X)$:

$$\sigma = \sqrt{Var(X)}$$

Variance of a linear function of a RV X

- For any pair of constants (a, b),

$$Var(aX + b) = a^2 Var(X)$$

- Easy to demonstrate :

$$\begin{aligned} Var(aX+b) &= E[(aX+b) - E(aX+b)]^2 \\ &= E[aX + b - (aE[X] + b)]^2 \\ &= E[aX - aE[X]]^2 \\ &= E[a^2(X - E[X])^2]] \\ &= a^2 E[(X-\mu)^2] \\ &= a^2 Var(X) \end{aligned}$$

- Thus for a linear function of X, we have :

$$SD(aX + b) = |a| \; SD(X)$$

- ('SD' = 'standard deviation')

Bernoulli RV

- A *Bernoulli* RV takes on only values 0 and 1
- The pmf for a Bernoulli RV is :

x	0	1
$p(x)$	$(1-p)$	p

- Used for modeling situations where there are 2 possible outcomes on a trial : heads/tails (coin tossing) ; yes/no ; success/failure ; *etc.*
- For a Bernoulli RV X :

$$E(X) = 0 \times (1-p) + 1 \times p = \mathbf{p};$$

$$Var(X) = E[X^2] - (E[X])^2 = [0^2 \times (1-p) + 1^2 \times p] - p^2$$
$$= p - p^2 = \mathbf{p(1\text{-}p)}$$

Probability mass function for a binomial RV

- Easy to derive using fundamental principles (as we have already done !)
- There are **4 conditions** to satisfy :
  - a *fixed* (not random) number n of trials
  - result for each trial : *either 1, or 0*
  - *the same probability* p for each trial of obtaining 1
  - the trials are *independent*
- Thus, if $X \sim Bin(n, p)$, the pmf is :

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \; ; \quad \underline{\underline{x = 0, 1, \ldots, n}}$$

Expected value and variance of a binomial RV

- I toss a fair coin 10 times (independently) How many times do you expect the coin to land Heads *??*
- I toss a fair 6-sided die 60 times. How many times do you expect the number '1' to come up *??*
- The *expected value* of $X \sim Bin(n, p)$ is $E[X] = np$ (corresponding to our intuition)
- The *variance* is $Var(X) = np(1-p)$ (which is less intuitive!!)

Proof : $E[X]$ (DON'T NEED TO KNOW THIS)

$$
\begin{aligned}
E[X] &= \sum_{k=0}^{n} kP(X=k) = \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k (1-p)^{n-k} \\
&= \sum_{k=\mathbf{1}}^{n} k \cdot \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\
&= \sum_{k=1}^{n} k \cdot \frac{n \cdot (n-1)!}{k \cdot (k-1)!(n-k)!} p \cdot p^{k-1} (1-p)^{n-k} \\
&= np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-k)!} p^{k-1} (1-p)^{n-k}
\end{aligned}
$$

Re-label the indices of the sum : $m = n-1; i = k-1$. This does not change the value of the sum, but it is easier to see that *we are in fact summing over all possible values of a binomial RV :*

$$
= \quad np \sum_{i=0}^{m} \frac{m!}{i!(m-i)!} p^i (1-p)^{m-i} = np \cdot 1 = \mathbf{np}
$$

(The derivation for the variance is similar.)

Discrete RV : another example

Example 4.3 The pmf of X is : $p(i) = c\lambda^i / i!$, $i = 0, 1, 2, \ldots$, where λ is a positive real value. Find :

(a) $P(X = 0)$ **(b)** $P(X > 2)$

Solution

- First, we find the value c ; since $\sum_{i=0}^{\infty} p(i) = 1$:

$$c \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = 1$$

- Recalling that $\sum_{i=0}^{\infty} \frac{x^i}{i!} = e^x$, we have $ce^{\lambda} = 1 \implies \boxed{c = e^{-\lambda}}$
- This RV is a *Poisson RV*

Solution, cont.

Thus :

(a) $P(X = 0) = e^{-\lambda}\lambda^0/0! = \boxed{e^{-\lambda}}$ [substitution $(i = 0)$ in $p(i)$]

(b)

$$\begin{aligned} P(X > 2) &= 1 - P(X \leq 2) && \text{[prob. complementary ev.]} \\ &= 1 - P(X = 0) - P(X = 1) - P(X = 2) && \text{[ME events]} \\ &= \boxed{1 - e^{-\lambda} - \lambda e^{-\lambda} - \frac{\lambda^2 e^{-\lambda}}{2}} && \text{[substitution for } p(i)\text{]} \end{aligned}$$

BREAK

Approximation of the binomial dist by that of Poisson

- We can use a Poisson RV *to approximate a binomial RV* with parameters (n, p) in the case where n is large and with parameters (n, p) in the case where :

  n is large and *p is small* but *np is of a moderate size*

- Proof : Let $X \sim Bin(n, p)$ and $\lambda = np$; then

$$P(X = i) = \frac{n!}{(n-i)!i!} \, p^i (1-p)^{n-i} = \frac{n!}{(n-i)!i!} \left(\frac{\lambda}{n}\right)^i \left(1 - \frac{\lambda}{n}\right)^{n-i}$$

$$= \frac{n(n-1)\cdots(n-i+1)}{n^i} \cdot \frac{\lambda^i}{i!} \cdot \frac{(1-\lambda/n)^n}{(1-\lambda/n)^i}$$

- For n large and λ moderate,

$$\left(1 - \frac{\lambda}{n}\right)^n \approx e^{-\lambda} \qquad \frac{n(n-1)\cdots(n-i+1)}{n^i} \approx 1 \qquad \left(1 - \frac{\lambda}{n}\right)^i \approx 1$$

- $\Rightarrow$ Thus $\boxed{P(X = i) \approx e^{-\lambda}\lambda^i / i! \;; \quad x = 0, 1, 2, \ldots}$

Applications of the Poisson distribution

- Some examples of RV $\sim$ Poisson :
  - number of typos per page of a book
  - number of false telephone numbers dialed in a day
  - number of clients entering a particular post office in a day
  - number of α particles emitted by a radioactive substance during a certain time lapse
  - number of bacterial colonies multiplying in a Petri dish in a nutritious medium
  - number of mutants resulting from an experiment
- In each of these examples (and in many others) the RV is always distributed as *approximately Poisson*, following a binomial distribution with parameter n large and parameter p small (even if we don't know the 'true' value n)

Example

Example 4.4 Suppose that the number of typographical errors on a single page of a particular book has a Poisson distribution with parameter $\lambda = \frac{1}{2}$. Calculate the probability that there is at least one error on page 27 ...

Another example

Example 4.5 Consider an experiment that consists of counting the number of α-particles given off in a 1-second interval by 1 gram of radioactive material. on average, the number of α-particles emitted is 3.2.
Give a good approximation to the probability that at most two α-particles will appear ...

Solution Let's represent the gram of radioactive material as a collection of a large number n of atoms. Each one can disintegrate, this having a probability of $3.2/n$ during the measured duration and giving off an α-particle during this time.

We can then say that X = number of α-particles emitted has an approximate Poisson distribution with parameter $\lambda =$ ____

Then $P(X \leq 2) =$ ...

Expected value and Variance for Poisson RVs : Intuition

- Recall that a Poisson RV is an approximation of a binomial RV with parameters n and p when n large, p is small and $\lambda = np$
- Let $X \sim Bin(n, p)$ (n large, p small, np moderate) :
  - $\lambda = np = E[X]$
  - $Var(X) = np(1-p) = \lambda(1-p) \approx \lambda$ (if p is small)
- So it would seem that **BOTH** the expected value and variance should be **EQUAL to the parameter λ**
- We can *verify this intuition* with a calculation (which **will not be examined**)
- **[if it's interesting for you, see these calculations on moodle]**

Expected value of Poisson RVs : **Calculation**

$$
\begin{aligned}
E[X] &= \sum_{i=0}^{\infty} \frac{ie^{-\lambda}\lambda^{i}}{i!} && \underline{\hspace{8cm}} \\
&= \lambda \sum_{i=1}^{\infty} \frac{e^{-\lambda}\lambda^{i-1}}{(i-1)!} && \underline{\hspace{8cm}} \\
&= \lambda e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^{j}}{j!} && (\text{posing } j = i-1) \\
&= \lambda && \left(\text{since } \sum_{j=0}^{\infty} \frac{\lambda^{j}}{j!} = \underline{\hspace{1.5cm}}\right)
\end{aligned}
$$

Variance of Poisson RVs : Calculation

$$
\begin{aligned}
E[X^2] &= \sum_{i=0}^{\infty} \frac{i^2 e^{-\lambda}\lambda^i}{i!} \\
&= \lambda \sum_{i=1}^{\infty} \frac{i e^{-\lambda}\lambda^{i-1}}{(i-1)!} \\
&= \lambda \sum_{j=0}^{\infty} \frac{(j+1) e^{-\lambda}\lambda^{j}}{j!} \qquad (\text{posing } j = i-1) \\
&= \lambda \left[\sum_{j=0}^{\infty} \frac{j e^{-\lambda}\lambda^{j}}{j!} + \sum_{j=0}^{\infty} \frac{e^{-\lambda}\lambda^{j}}{j!} \right] \\
&= \lambda(\lambda + 1) \qquad [\text{substitution } E(X), \sum_i p(i) = 1]
\end{aligned}
$$

Example

Example 4.5 In a microbial mutagenesis assay, a plate of bacteria is exposed to a test compound, and the number of mutants is counted after incubation. Suppose that in the assay of a particular compound, the number of mutants has a Poisson distribution with $\lambda = 9$.

Calculate the probability that an assay of the compound will produce :

(a) 0 mutants

(b) more than 3 mutants

Another example

Example 4.6 Let Y be the number of assays with at least one mutant in 5 independent assays of the compound from Example 4.5.

(a) What is a reasonable *probability model* for Y *??* Explain.

(b) What are the *parameter values* for the model *??*

(c) What is the probability that there is at least 1 mutant in (exactly) 2 of the 5 assays *??* [Hint : use the 4 steps...]