PROBABILITY THEORY
MM 4

MM 4: Special probability distributions

Topics:

Discrete distributions:

Bernoulli, binomial, geometric, Poisson

Quizz

Suppose X has mean 2 and variance 3

- Compute Var(3X)
  - 3
  - 9
  - 27
- Compute Var(3X+4)
  - 27
  - 31
  - 43
- Compute E[X^2]
  - 4
  - 7
- Compute Var(X^2)
  - 9
  - Can not be computed from infomration given

Quizz

- Can the following be negative:
  - A r.v. X
  - E[X]
  - E[X^2]
  - Var(X)
  - Cov(X,Y)

Quiz

- If you have a pdf (pmf), can you find a cdf?
- Is the opposite true?

The nth moment of r.v. X

- The mean value is sometimes called the first moment of X.
- Definition. The nth moment of X is defined by

$$\boxed{E[X^n] = \sum_i x_i^n p(x_i)}$$

$$\boxed{E[X^n] = \int_{-\infty}^{\infty} x^n f(x) dx}$$

Moment generating function

- Definition. The moment generating function of a r.v. X is defined for all values t by

$$\varphi(t) = E[e^{tX}] = \sum_i e^{tx_i} p(x_i)$$

$$\varphi(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f(x) dx$$

- This function is called moment generating function, because all moments of X can be obtained by successively differentiating it.
- There is a one-to-one correspondence between the moment generating function and the distribution function of a r.v.: the mgf uniquely determines the distribution and vice versa.

How to obtain E[X] and Var(X) from moment generating function?

$$\varphi'(t) = \frac{d}{dt} E[e^{tX}] = E\left[\frac{d}{dt}(e^{tX})\right] =$$

$$= E[X e^{tX}]$$

$$\varphi'(0) = E[X]$$

$$\varphi''(t) = \frac{d}{dt} E[X e^{tX}] = E\left[\frac{d}{dt}(X e^{tX})\right] =$$

$$= E[X^2 e^{tX}]$$

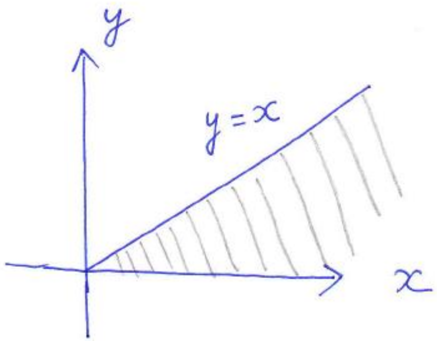
$$\varphi''(0) = E[X^2]$$

$$Var(X) = \varphi''(0) - \varphi'(0)^2$$

Example: covariance

joint pdf for X and Y:

$$f_{X,Y}(x,y) = \begin{cases} c e^{-x} e^{-y}, & 0 \le y < x < \infty \\ 0, & \text{otherwise} \end{cases}$$



The constant c is found from the normalization condition:

$$1 = \int_0^{+\infty} \int_0^{x} c e^{-x} e^{-y} \, dy \, dx = \frac{c}{2} \quad \Rightarrow \quad c = 2$$

Example: covariance (ctnd)

The marginal pdfs:

$$f_X(x) = \int_0^x 2e^{-x}e^{-y}\,dy = 2e^{-x}(1-e^{-x}), \quad x \in [0,\infty)$$

$$f_Y(y) = \int_0^\infty 2e^{-x}e^{-y}\,dx = 2e^{-2y}, \quad y \in [0,\infty)$$

$$E[X] = \frac{3}{2}, \quad Var(X) = \frac{5}{4}$$

$$E[Y] = \frac{1}{2}, \quad Var(Y) = \frac{1}{4}$$

Example: covariance (ctnd)

$$E[XY] = \int_0^\infty \int_0^x xy \cdot 2e^{-x} e^{-y} \, dy \, dx =$$

$$= \int_0^\infty 2x e^{-x} (1 - e^{-x} - x e^{-x}) \, dx = 1$$

$$Cov(X,Y) = E[XY] - E[X]E[Y] = 1 - \frac{3}{4} = \frac{1}{4}$$

$$Corr(X,Y) = \frac{1/4}{\sqrt{5/4}\,\sqrt{1/4}} = \frac{1}{\sqrt{5}}$$

What should we learn today?

- A number of r.vs. arises in many diverse, unrelated applications → learn them
- What are the main facts about these distributions:
  - Their cdf and pdf
  - Mean and variance
  - Moment generating function
- In which situations the special distributions arrise and how are they interrelated?

Bernoulli r.v.

- An experiment with 2 possible outcomes (success or failure) is called a Bernoulli trial.
- The indicator of event A is called the Bernoulli r.v., since it describes outcome of a Bernoulli trial.

$$E[I] = p \qquad\qquad Var(I) = p(1-p)$$

- Every Bernoulli trial, regardless of the definition of A, is equivalent to the tossing of a biased coin.

Example: indicator of an event

- Example: a r.v. I is an indicator of the event A

- *The expected value of the indicator of an event is equal to the probability of the event*

$$I = \begin{cases} 1, & \text{if } A \text{ occurs} \\ 0, & \text{if } A \text{ does not occur} \end{cases}$$

$$E[I] = 1 \cdot P(A) + 0 \cdot P(A^c) = P(A)$$

Example: indicator of an event

$$I = \begin{cases} 1, & \text{if event } A \text{ occurs} \\ 0, & \text{if event } A \text{ does not occur} \end{cases}$$

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

$$I^2 = I$$

$$= E[I](1 - E[I]) = P(A)(1 - P(A)) = P(A)P(A^c)$$

Binomial r.v.

- Let X be a r.v. representing the number of times a certain event A occurs in n trials (number of success in n trials). Denote p probability of success.
- Then X is said to be a binomial r.v. with parameters n and p

$$X \sim b(n, p)$$

- Pmf

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

$$E[X] = np \qquad Var(X) = np(1-p)$$

Binomial r.v.: mean and variance

$X = I_1 + I_2 + .. + I_n$

where I_i is a Bernoulli r.v.

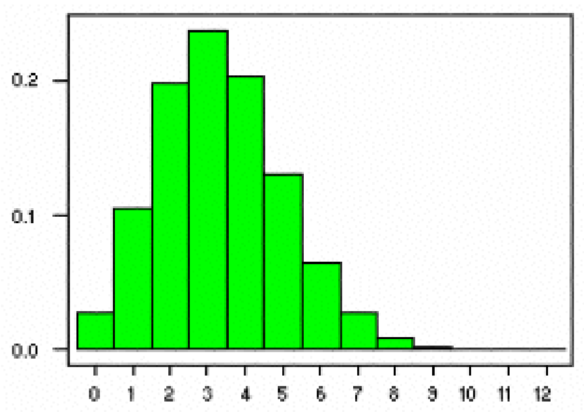
$E[X] = E[I_1 + I_2 + .. + I_n] = E[I_1] + E[X_2] + ... + E[X_n] =$

$= \underbrace{p + p + .. + p}_{n \text{ times}} = np$

$Var(X) = Var(I_1 + ... + I_n) = Var(I_1) + .. + Var(I_n) =$

$= \underbrace{p(1-p) + ... + p(1-p)}_{n \text{ times}} = np(1-p)$

Binomial r.v.



Parameters n=20, p=1/6

Computing the binomial distribution function

$$P\{X \le i\} = \sum_{k=0}^{i} P\{X=k\} = \sum_{k=0}^{i} \binom{n}{k} p^k (1-p)^{n-k}$$

- To optimize the computational process, we can utilize the following relationship between P{X=k+1} and P{X=k}:

$$P\{X = k+1\} = \frac{p}{1-p}\frac{n-k}{k+1}P\{X=k\}$$

- Now starting from P(X=0) we can recursively find all other P(X=k)

Example: n-component system

- A system consists of n components. Each component functions independently with probability p.
- Question: for which values of p is a 5-component system more likely to operate than a 3-component system?

For 5-component system:

$$P_1 = P\{X \geq 3\} = P\{X=3\} + P\{X=4\} + P\{X=5\} =$$

$$= 1 - P\{X<3\} = 1 - P\{X=0\} - P\{X=1\} - P\{X=2\} =$$

$$= \binom{5}{3} p^3 (1-p)^2 + \binom{5}{4} p^4 (1-p) + \binom{5}{5} p^5 (1-p)^0$$

Example (cntd)

For 3-component system:

$$P_2 = P\{Y \geq 2\} = P\{Y=2\} + P\{Y=3\} =$$

$$= \binom{3}{2} p^2(1-p) + \binom{3}{3} p^3(1-p)^0$$

We would like to choose p such that $P_1 \geq P_2$

$$\Rightarrow \quad 10p^3(1-p)^2 + 5p^4(1-p) + p^5 \geq 3p^2(1-p) + p^3$$

$$3(1-p)^2(2p-1) \geq 0$$

$$p \geq \frac{1}{2}$$

Why is it called binomial?

N / k	0	1	2	3	4
1	p	q			
2	p^2	$2pq$	q^2		
3	p^3	$3p^2q$	$3pq^2$	q^3	
4	p^4	$4p^3q$	$6p^2q^2$	$4pq^3$	q^4

Coefficients are the same as in the binomial expansion $(p+q)^n$

The geometric r.v.

- Let a r.v. M be a number of independent Bernoulli trials until the first occurence of a success. M is called a geometric r.v.
- Pmf

$$P\{M = k\} = (1-p)^{k-1}p$$

$$E[M] = \frac{1}{p} \qquad Var(M) = \frac{1-p}{p^2}$$

- Example: applications where we are interested in the time that elapses between the occurence of events in a sequence of independent experiments

Geometric r.v.: mean and variance

$$P\{M = k\} = (1-p)^{k-1}p$$

$$E[M] = \sum_{k=1}^{+\infty} k\, p\, (1-p)^{k-1}$$

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$$

$$\sum_{k=0}^{+\infty} k\, x^{k-1} = \frac{1}{(1-x)^2}$$

$$x = 1-p$$

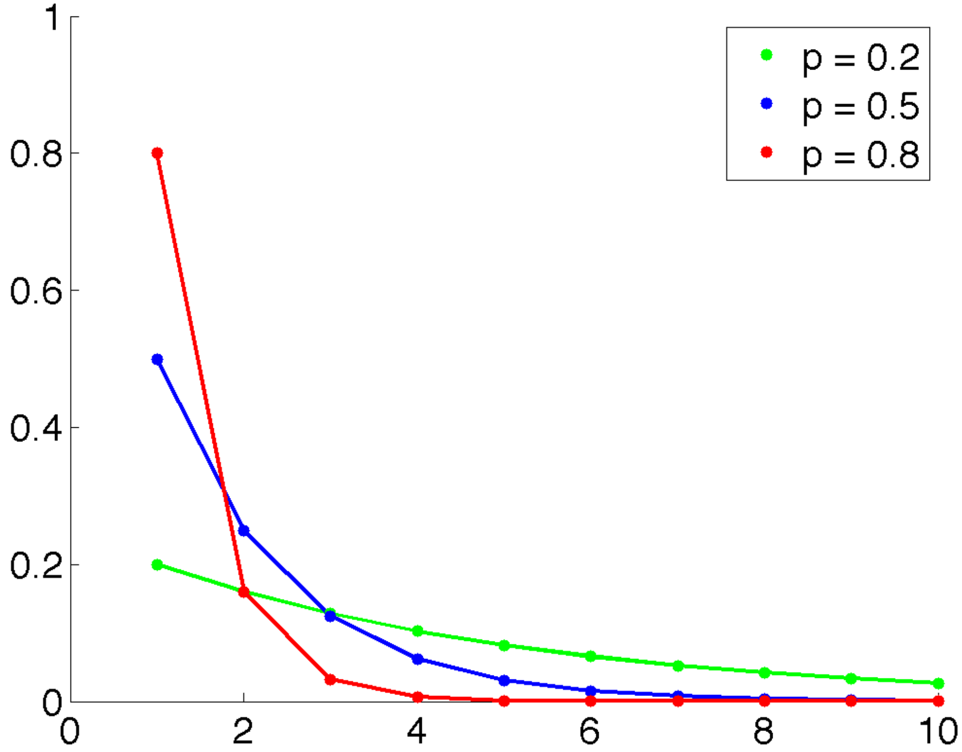
$$E[M] = p \cdot \frac{1}{(1-(1-p))^2} = \frac{p}{p^2} = \frac{1}{p}$$

M	Probability
1	p
2	$(1-p)p$
3	$(1-p)^2 p$
⋮	
k	$(1-p)^{k-1}p$
⋮	

M^2	Pr
1	p
2^2	$(1-p)p$
3^2	$(1-p)^2 p$
⋮	
k^2	$(1-p)^{k-1}p$

$$E[M] = \frac{1}{p} \qquad Var(M) = \frac{1-p}{p^2}$$

The geometric r.v.



The hypergeometric r.v.

- Suppose we have objects of two types: type 1 N objects (acceptable components) and type 2 M objects (defective components). A sample of n objects is randomly chosen without replacement.
- A r.v. X represents a number of type 1 objects in the selection → hypergeometric with parameters (N, M, n)
- Pmf

$$P\{X=i\} = \frac{\binom{N}{i}\cdot\binom{M}{n-i}}{\binom{N+M}{n}}$$

$$E[X] = \frac{nN}{N+M} \qquad Var(X) = np(1-p)[1-\frac{n-1}{N+M-1}]$$

Derivation of formulas for mean and variance of hypergeometric r.v.

$$X = I_1 + I_2 + \ldots + I_n$$

$$P\{I_k = 1\} = \frac{N}{N+M} = p$$

$$E[X] = \sum_{k=1}^{n} E[I_k] = np = \frac{nN}{N+M}$$

I_k are not independent $\Rightarrow$

$$Var(X) = \sum_{k=1}^{n} Var(I_k) + 2 \sum\sum_{1 \le k < \ell \le n} Cov(I_k, I_\ell)$$

$$Var(I_k) = p(1-p)$$

$$Cov(I_k, I_\ell) = E[I_k I_\ell] - E[I_k]E[I_\ell]$$

$$E[I_k I_\ell] = P\{I_k I_\ell = 1\} = P\{I_k = 1, I_\ell = 1\} =$$

$$= \frac{N}{N+M} \cdot \frac{N-1}{N+M-1}$$

Poisson r.v.

- A r.v. taking on one of values 0, 1, 2, … is said to be a Poisson r.v. with parameter λ (λ>0), if pmf is given by

$$P\{X = i\} = e^{-\lambda} \frac{\lambda^i}{i!}$$

- Moment generating function:

$$\phi(t) = E[e^{tX}] = e^{\lambda(e^t - 1)}$$

$$E[X] = \lambda \qquad Var(X) = \lambda$$

Examples

- Examples of events that can be modelled by Poisson distribution:
  - Number of typos (misprints) on a page of a book
  - Number of phone calls at a call center per minute
  - Number of times a web server is accessed per minute
  - Number of pine trees per unit area in a mized forest
  - Number of stars in a given volume of space
  - …

When is the Poisson distribution an appropriate model?

- The occurrence of one event does not affect the probability that a second event will occur. That is, events occur independently.
- The rate at which events occur is constant. The rate cannot be higher in some intervals and lower in other intervals.
- The probability of an event in a small interval is proportional to the length of the interval.

Poisson r.v.

Lets verify that it is truly a pmf:

$$\sum_{i=0}^{\infty} P\{X=i\} = \sum_{i=0}^{\infty} e^{-\lambda}\frac{\lambda^i}{i!} = e^{-\lambda}\sum_{i=0}^{\infty}\frac{\lambda^i}{i!} =$$

$$= e^{-\lambda}\cdot e^{\lambda} = 1$$

Moment generation function:

$$\varphi(t) = E[e^{tX}] = \sum_{i=0}^{\infty} e^{ti} e^{-\lambda}\frac{\lambda^i}{i!} = e^{-\lambda}\sum\frac{(\lambda e^t)^i}{i!} =$$

$$= e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t-1)}$$

$$\varphi'(t) = \lambda e^t e^{\lambda(e^t-1)} \qquad \varphi'(0) = \lambda\cdot 1\cdot e^{\lambda(1-1)} = \lambda$$

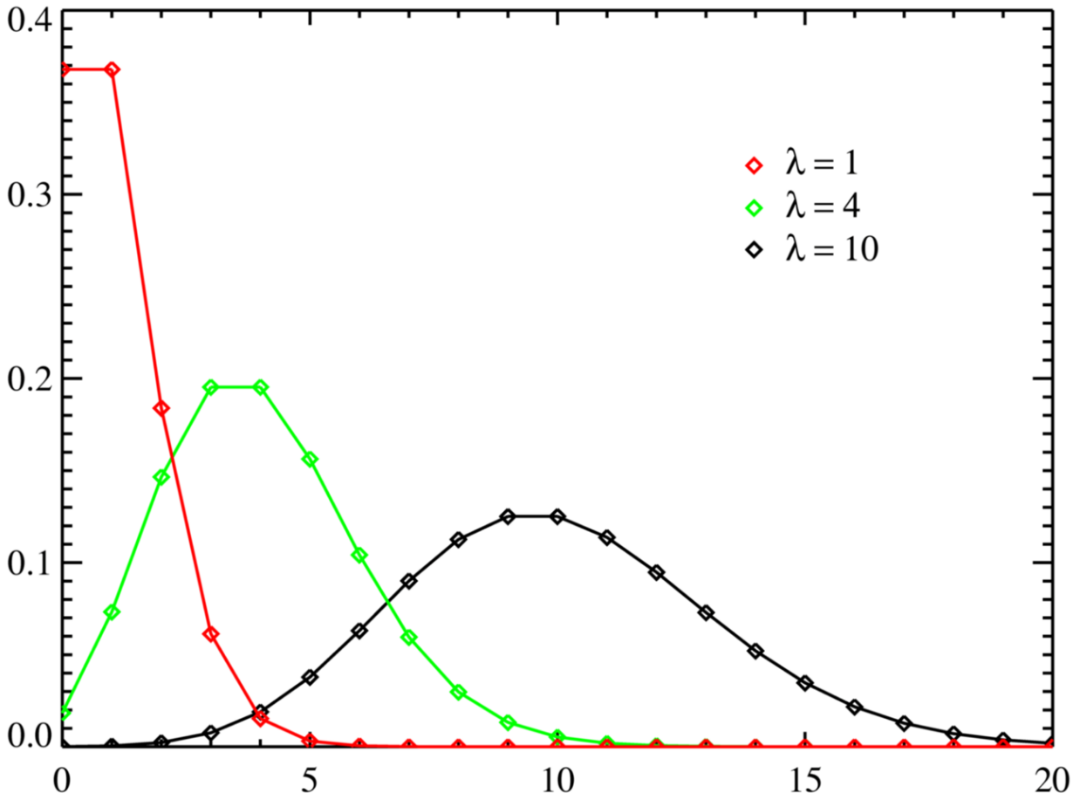
$$\Rightarrow E[X] = \lambda$$

$$\varphi''(t) = (\lambda e^t)^2 e^{\lambda(e^t-1)} + \lambda e^t e^{\lambda(e^t-1)}$$

$$\Rightarrow E[X^2] = \varphi''(0) = \lambda^2+\lambda$$

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

Poisson r.v.



Approximation for a binomial r.v.

- One of the applications of the Poisson probabilities is to approximate the binomial probabilities – when n is large and p is small.

- $\lambda = np$

Let $X \sim b(n,p)$. We put $\lambda = np$

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

$$= \underbrace{\frac{n(n-1)\cdot\ldots\cdot(n-i+1)}{n^i}}_{\xrightarrow{n\to\infty} 1} \cdot \frac{\lambda^i}{i!} \; \frac{\overbrace{\left(1-\frac{\lambda}{n}\right)^n}^{\xrightarrow{n\to\infty} e^{-\lambda}}}{\underbrace{\left(1-\frac{\lambda}{n}\right)^i}_{\xrightarrow{n\to\infty} 1}} \xrightarrow[n\to\infty]{} e^{-\lambda} \cdot \frac{\lambda^i}{i!}$$

Properties of Poisson distribution

- The probability of at least one occurrence of the event in a given time interval is proportional to the length of the interval.
- The probability of two or more occurrences of the event in a very small time interval is negligible.
- The numbers of occurrences of the event in disjoint time intervals are mutually independent.

Example

- The BER of a communication channel is 10^-3. What is a probability that a block of 1000 bits has five or more errors?

$X \sim b(n, p) \quad n = 1000, \; p = 10^{-3}$

Poisson approximation: $\lambda = np = 1$

$$P\{X \geq 5\} = 1 - P\{X < 5\} = 1 - \sum_{k=0}^{4} \frac{\lambda^k}{k!} e^{-\lambda} =$$

$$= 1 - e^{-1}\left(1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!}\right) \simeq 0.0036$$

Properties of a Poisson r.v.

- Reproductive property
- Proposition.

1) The sum of independent Poisson r.vs. with parameters λ_1 and λ_2 is also a Poisson r.v. with parameter $\lambda_1 + \lambda_2$

2) If a Poisson event having mean λ can be independently classified as being of type 1, 2, … r with respective probabilities

$$p_1, p_2, \ldots, p_r (\sum p_i = 1)$$

then type i events are independent Poisson r.v. with means λp_1 , $\lambda p_2 \ldots \lambda p_r$