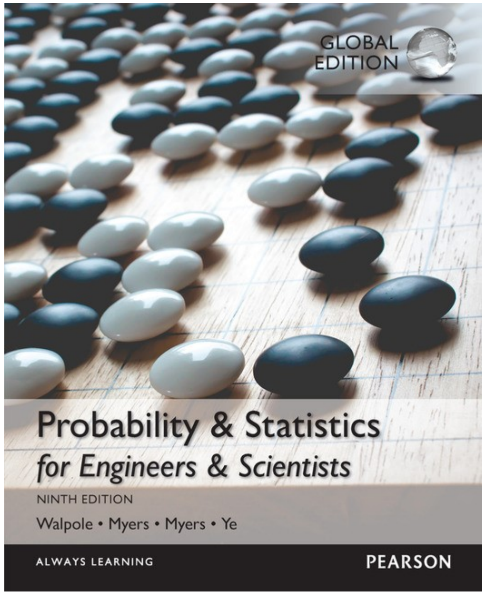
Chapter 5

Some Discrete Probability Distributions– part 2

School of Computing, Gachon Univ.
Joon Yoo



Wireless Intelligent Networking LAB

- Question 1
  - Toss a dice repeatedly. What's the probability that we observe a '3' for the first time on the 10-th toss?

- Question 2
  - In a Korean Baseball championship series, the team who wins four games out of seven will be the winner. Suppose that team S has probability 0.55 of wining over D, and both teams S and D face each other in the championship games.
    - 1. What is the probability that team S will win the series in 5 games?
    - 2. What is the probability that team D will win the series?

5.4 Negative Binomial and Geometric Distributions

Negative Binomial Experiment

- Let us consider an experiment where the properties are the same as those listed for a binomial experiment, with the exception that the trial will be repeated until a fixed number of successes occur.

- Therefore, instead of finding the probability of x successes in n trials (where n is fixed), we are now interested in the probability that the k-th success occurs on the x-th trial. Experiments of this kind are called <u>negative binomial experiments</u>.

Negative Binomial Experiment

- As an illustration, consider the use of a drug that is known to be effective in 60% of the cases where it is used. The drug will be considered a success if it is effective in bringing some degree of relief to the patient. We are interested in finding the probability that the **fifth patient** to experience relief is the **seventh patient** to receive the drug during a given week. Designating a success by **S** and a failure by **F**, a possible order of achieving the desired result is SFSSSF**S**, which occurs with probability

$$(0.6)(0.4)(0.6)(0.6)(0.6)(0.4)(0.6) = (0.6)^5(0.4)^2.$$

- We could list all possible orders by rearranging the F's and S's except for the last outcome, which must be the fifth success. The total number of possible orders is equal to the number of partitions of the first six trials into two groups with 2 failures assigned to the one group and 4 successes assigned to the other group. This can be done in $\binom{6}{4} = 15$ mutually exclusive ways. Hence, if X represents the outcome on which the fifth success occurs, then

$$P(X = 7) = \binom{6}{4}(0.6)^5(0.4)^2 = 0.1866.$$

- # Negative Binomial Experiment
  - If repeated independent trials can result in a success with probability $\boldsymbol{p}$ and a failure with probability $\boldsymbol{q} = 1 - p$, then the probability distribution of the random variable X, the number of the trial on which the ***k*-th** success occurs, is denoted as

$$b^*(x; k, p) \text{ or } nb(x; k, p)$$

    - X = the number of the trials on which the k-th success occurs if and only if get exactly k – 1 successes by time x – 1, and then the k-th success at time x

$$P(X = x) = \left[\binom{x-1}{k-1} p^{k-1} q^{x-k}\right] p, \quad x = k, k+1, k+2, \ldots.$$
$$= \binom{x-1}{k-1} p^k q^{x-k}, \quad x = k, k+1, k+2, \ldots.$$

$$b^*(x; k, p) = \binom{x-1}{k-1} p^k q^{x-k}, \quad x = k, k+1, k+2, \ldots.$$

- **Negative Binomial Experiment**
  - If repeated independent trials can result in a success with probability p and a failure with probability $q = 1 - p$, then the probability distribution of the random variable X, the number of the trial on which the ***k*-th** success occurs, is

$$P(X=x) = nb(x;k,p) = \binom{x-1}{k-1} p^k q^{x-k}, \quad x = k, k+1, k+2, \ldots.$$

$$E(X) = \frac{k}{p} \quad \text{and} \quad \sigma^2 = \mathrm{Var}(X) = \frac{k(1-p)}{p^2}$$

- Example
  - Find the probability that a person flipping a fair coin gets the second head on the sixth flip.

- Solution
  - Let X be the number of flips needed to get the 2nd head.

$$X \sim \mathrm{NB}(2, 0.5), \quad nb(6; 2, 0.5) = P(X = 6) = \binom{5}{1} (0.5)^2 (0.5)^4.$$

$$\boxed{b^*(x; k, p) = \binom{x-1}{k-1} p^k q^{x-k}, \quad x = k, k+1, k+2, \ldots.}$$

- Example 5.14

  - In an NBA (National Basketball Association) championship series, the team who wins four games out of seven will be the winner. Suppose that team A has probability 0.55 of wining over B and both teams A and B face each other in the championship games.
    - 1. What is the probability that team A will win the series in six games?
    - 2. What is the probability that team A will win the series?

Wireless Intelligent Networking LAB

- Solution
  - Let X represent the number of games that team A needs to play to win the championship.
  - X~b*(4, 0.55)   // = nb(4, 0.55)
  - 1.

$$b^*(x;k,p) = \binom{x-1}{k-1} p^k q^{x-k}, \quad x = k, k+1, k+2, \ldots.$$

$$\begin{aligned} P(X=6) &= nb(6;4,0.55) \\ &= \binom{5}{3}(0.055)^4(1-0.55)^{6-4} \\ &= 0.1853. \end{aligned}$$

  - 2.

$$\begin{aligned} & P(\text{team A wins the championship series}) \\ =& P(X=4)+P(X=5)+P(X=6)+P(X=7) \\ =& nb(4;4,0.55)+nb(5;4,0.55)+nb(6;4,0.55)+nb(7;4,0.55) \\ =& 0.6083. \end{aligned}$$

Wireless Intelligent Networking LAB

Geometric Distribution

- If we consider the special case of the negative binomial distribution where **k = 1**, we have a probability distribution for the **number of trials required for a single success**.
- An example would be the tossing of a coin until a head occurs. We might be interested in the probability that the first head occurs on the fourth toss. The negative binomial distribution reduces to the form

$$\boxed{b^*(x;k,p) = \binom{x-1}{k-1} p^k q^{x-k}, \quad x = k, k+1, k+2, \ldots.}$$

$$b^*(x;1,p) = pq^{x-1}, \quad x = 1, 2, 3, \ldots.$$

- Since the successive terms constitute a geometric progression, it is customary to refer to this special case as the **geometric distribution** and denote its values by $g\,(x;\,p)$.

Geometric Distribution

- If repeated independent trials can result in a success with probability *p* and a failure with probability $q = 1 - p$, then the probability distribution of the random variable X, **the number of the trial on which the first success occurs**, is

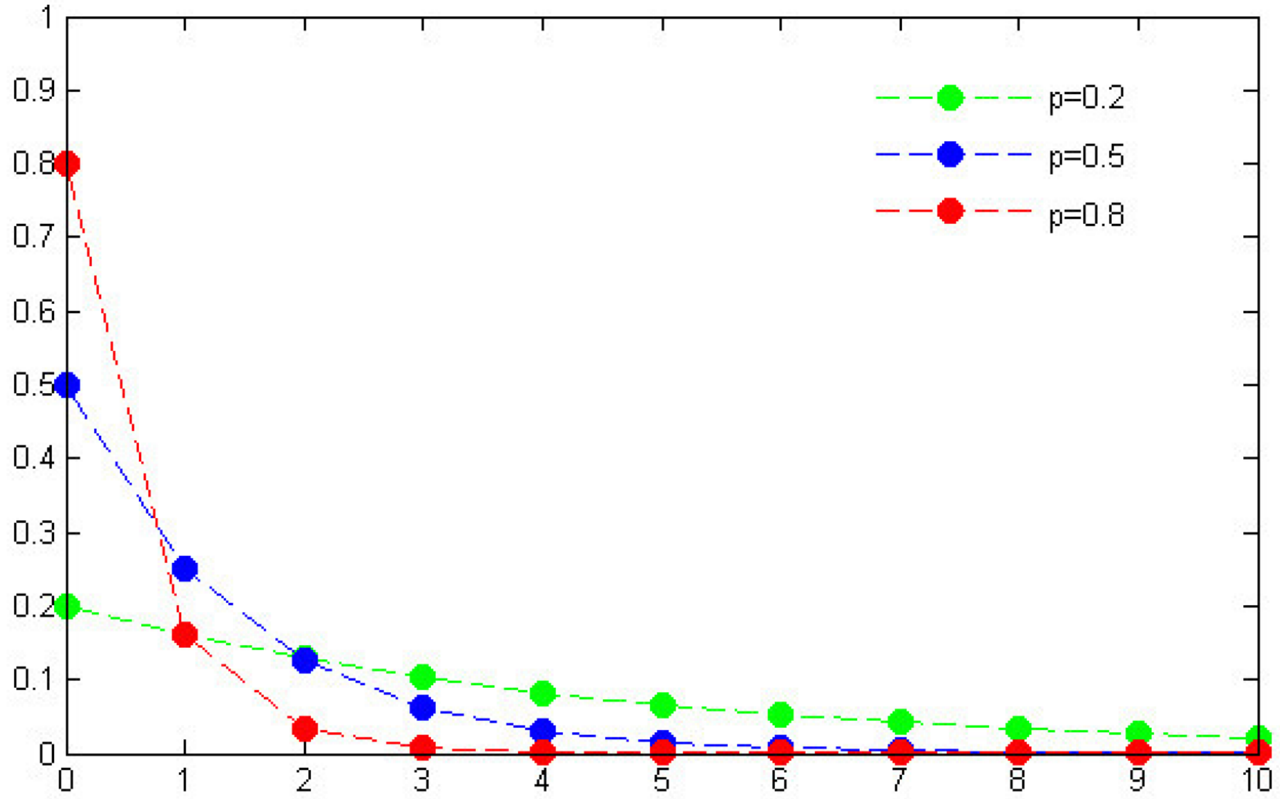
$$P(X = x) = g(x; p) = pq^{x-1}, \quad x = 1, 2, 3, \ldots .$$

$$E(X) = \frac{1}{p} \quad \text{and} \quad \sigma^2 = \mathrm{Var}(X) = \frac{1-p}{p^2}$$

The mean and variance of a random variable following the geometric distribution are

$$\mu = \frac{1}{p} \text{ and } \sigma^2 = \frac{1-p}{p^2}.$$

*We do not give proof!



Probability density function for geometric random variables with different values of p.

- Example
  - Toss a dice repeatedly. What's the probability that we observe a '3' for the first time on the 8th toss?

- Solution
  - Let X be number of trials required to produce the first 3. So,

$$X \sim \text{Geo}(p = \frac{1}{6}).$$

  - Using the geometric distribution with X = 8 and p = 1/6, we have

$$P(X = 8) = g(8; \frac{1}{6}) = \left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^7.$$

$$\boxed{P(X = x) = g(x; p) = pq^{x-1},}$$

- Q: How many tosses would we expect to take?

$$E(X) = \frac{1}{p} = 6 \text{ tosses.}$$

$$\boxed{\mu = \frac{1}{p} \text{ and } \sigma^2 = \frac{1-p}{p^2}.}$$

Example 5.16

- At a "busy time," a telephone exchange is very near capacity, so callers have difficulty placing their calls. It may be of interest to know the number of attempts necessary in order to make a connection.
- Suppose that we let p = 0.05 be the probability of a connection during a busy time. What is the probability that 5 attempts are necessary for a successful call?

Using the geometric distribution with $x = 5$ and $p = 0.05$ yields

$$P(X = x) = g(5; 0.05) = (0.05)(0.95)^4 = 0.041.$$

- What is the mean and variance of number of calls necessary for a successful call?

$$\mu = \frac{1}{p} = 1/0.05 = 20$$

$$\sigma^2 = \frac{1-p}{p^2} = 0.95/0.05^2 = 380$$

Question.

- Q1. If you rolled 5 times a fair die and never got a 6, what is the probability that you get 6 on the 6^{th} roll ?

- Q2. If you rolled 50 times a fair die and never got a 6, what is the probability that you get 6 on the 51^{th} roll ?

- If you rolled *100* times a fair die and never got a 6, you don't have more chances on the 101-th roll to get a 6 than on the first roll.

오늘 3타수 무안타라면 이 선수 이번 타석
안타 나올 **확률**이 높다는 뜻이죠!

Memoryless Property

- Some lifetime problems exhibit a very important property called the "**memoryless property**".
- A random variable X has the property that "the future is independent of the past."
- In mathematics:

$$P(X > t+s \mid X > t) = P(X > s)$$

Wireless Intelligent Networking LAB

5.5 Poisson Process and Poisson Distribution

- Example 1
  - The mean number of accidents per month at a certain intersection is 1. What is the probability that in any given year 10 accidents will occur at this intersection?

- Example 2
  - The average number of customers arriving per hour at a certain restaurant is 7.
    Compute the probability that more than 10 customers will arrive in a 2-hour period.

Poisson Experiment

- Experiments yielding numerical values of a random variable X, the number of outcomes occurring during a given time interval t (or distance, area, or volume), are called **Poisson experiments**.

Poisson Distribution

- The probability associated with the number of occurrences in a given period of time is given by,

$$p(x;\lambda t)=\frac{e^{-\lambda t}(\lambda t)^{x}}{x!},\ x=0,1,2,\ldots$$

where,

λ = average number of outcomes per unit time or region

***t* =** time interval or region

(***mean = variance = μ = λt***)

$$p(x;\mu)=\frac{e^{-\mu}\mu^{x}}{x!},\ x=0,1,2,\ldots$$

$$\sum_{x=0}^{\infty}p(x;\mu)=\sum_{x=0}^{\infty}\frac{e^{-\mu}\mu^{x}}{x!}=e^{-\mu}\sum_{x=0}^{\infty}\frac{\mu^{x}}{x!}=e^{-\mu}e^{\mu}=1$$

Taylor series $e^{x}=1+x+\frac{1}{2!}x^{2}+\frac{1}{3!}x^{3}+\frac{1}{4!}x^{4}+\cdots$

Wireless Intelligent Networking LAB

Example

Example 5.17: During a laboratory experiment, the average number of radioactive particles passing through a counter in 1 millisecond is 4. What is the probability that 6 particles enter the counter in a given millisecond?

***Solution*:** Using the Poisson distribution with $x = 6$ and $\lambda t = 4$ and referring to Table A.2, we have

$$p(6;4) = \frac{e^{-4}4^6}{6!} = \sum_{x=0}^{6} p(x;4) - \sum_{x=0}^{5} p(x;4) = 0.8893 - 0.7851 = 0.1042.$$

Table A.2 Poisson Probability Sums $\sum_{x=0}^{r} p(x;\mu)$

	μ								
r	**1.0**	**1.5**	**2.0**	**2.5**	**3.0**	**3.5**	**4.0**	**4.5**	**5.0**
0	0.3679	0.2231	0.1353	0.0821	0.0498	0.0302	0.0183	0.0111	0.0067
1	0.7358	0.5578	0.4060	0.2873	0.1991	0.1359	0.0916	0.0611	0.0404
2	0.9197	0.8088	0.6767	0.5438	0.4232	0.3208	0.2381	0.1736	0.1247
3	0.9810	0.9344	0.8571	0.7576	0.6472	0.5366	0.4335	0.3423	0.2650
4	0.9963	0.9814	0.9473	0.8912	0.8153	0.7254	0.6288	0.5321	0.4405
5	0.9994	0.9955	0.9834	0.9580	0.9161	0.8576	0.7851	0.7029	0.6160
6	0.9999	0.9991	0.9955	0.9858	0.9665	0.9347	0.8893	0.8311	0.7622
7	1.0000	0.9998	0.9989	0.9958	0.9881	0.9733	0.9489	0.9134	0.8666
8		1.0000	0.9998	0.9989	0.9962	0.9901	0.9786	0.9597	0.9319
9			1.0000	0.9997	0.9989	0.9967	0.9919	0.9829	0.9682
10				0.9999	0.9997	0.9990	0.9972	0.9933	0.9863
11				1.0000	0.9999	0.9997	0.9991	0.9976	0.9945
12					1.0000	0.9999	0.9997	0.9992	0.9980
13						1.0000	0.9999	0.9997	0.9993
14							1.0000	0.9999	0.9998
15								1.0000	0.9999
16									1.0000

Service Call Example - Poisson Process

- Example
  - An average of 2.7 service calls per minute are received at a particular maintenance center. The calls correspond to a Poisson process. To determine personnel and equipment needs to maintain a desired level of service, the plant manager needs to be able to determine the probabilities associated with numbers of service calls.

Our Example: $\lambda = 2.7$ and $t = 1$ minute

- What is the probability that fewer than 2 calls will be received in any given minute?
- The probability that fewer than 2 calls will be received in any given minute is

$$P(X < 2) = P(X = 0) + P(X = 1)$$

$$P(X < 2) = \frac{e^{-2.7}(2.7)^0}{0!} + \frac{e^{-2.7}(2.7)^1}{1!}$$

- The mean and variance are both λt, so

$\mu = \lambda t =$ ____________________

WINL Wireless Intelligent Networking LAB

Service Call Example (Part 2)

- If more than 6 calls are received in a 3-minute period, an extra service technician will be needed to maintain the desired level of service. What is the probability of that happening?
  - *$\mu = \lambda t = (2.7)(3) = 8.1$*
  - *8.1 is not in the table; we must use basic equation*

$$p(X > 6) = 1 - p(X \le 6) = 1 - \left[\frac{e^{-8.1}(8.1)^0}{0!} + \frac{e^{-8.1}(8.1)^1}{1!} \cdots + \frac{e^{-8.1}(8.1)^6}{6!} \right]$$

- *Suppose $\lambda t = 8$; see table with μ* = 8 and r = 6

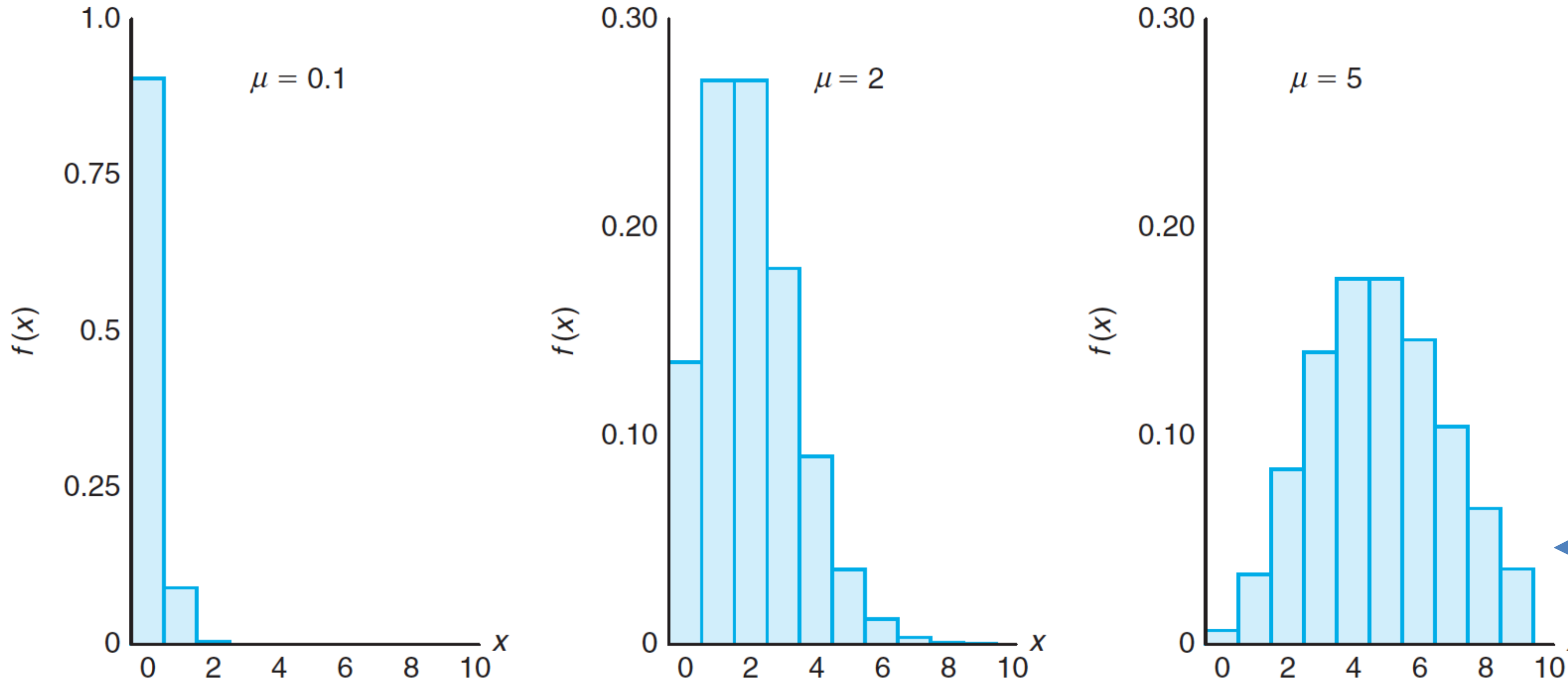
$P(X > 6) = 1 - P(X \le 6) = 1 - 0.3134 = 0.6866$

GWINL Wireless Intelligent Networking LAB

Table A.2 (continued) Poisson Probability Sums $\sum_{x=0}^{r} p(x;\mu)$

	μ								
r	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5
0	0.0041	0.0025	0.0015	0.0009	0.0006	0.0003	0.0002	0.0001	0.0001
1	0.0266	0.0174	0.0113	0.0073	0.0047	0.0030	0.0019	0.0012	0.0008
2	0.0884	0.0620	0.0430	0.0296	0.0203	0.0138	0.0093	0.0062	0.0042
3	0.2017	0.1512	0.1118	0.0818	0.0591	0.0424	0.0301	0.0212	0.0149
4	0.3575	0.2851	0.2237	0.1730	0.1321	0.0996	0.0744	0.0550	0.0403
5	0.5289	0.4457	0.3690	0.3007	0.2414	0.1912	0.1496	0.1157	0.0885
6	0.6860	0.6063	0.5265	0.4497	0.3782	0.3134	0.2562	0.2068	0.1649
7	0.8095	0.7440	0.6728	0.5987	0.5246	0.4530	0.3856	0.3239	0.2687
8	0.8944	0.8472	0.7916	0.7291	0.6620	0.5925	0.5231	0.4557	0.3918
9	0.9462	0.9161	0.8774	0.8305	0.7764	0.7166	0.6530	0.5874	0.5218
10	0.9747	0.9574	0.9332	0.9015	0.8622	0.8159	0.7634	0.7060	0.6453
11	0.9890	0.9799	0.9661	0.9467	0.9208	0.8881	0.8487	0.8030	0.7520
12	0.9955	0.9912	0.9840	0.9730	0.9573	0.9362	0.9091	0.8758	0.8364
13	0.9983	0.9964	0.9929	0.9872	0.9784	0.9658	0.9486	0.9261	0.8981
14	0.9994	0.9986	0.9970	0.9943	0.9897	0.9827	0.9726	0.9585	0.9400
15	0.9998	0.9995	0.9988	0.9976	0.9954	0.9918	0.9862	0.9780	0.9665
16	0.9999	0.9998	0.9996	0.9990	0.9980	0.9963	0.9934	0.9889	0.9823
17	1.0000	0.9999	0.9998	0.9996	0.9992	0.9984	0.9970	0.9947	0.9911
18		1.0000	0.9999	0.9999	0.9997	0.9993	0.9987	0.9976	0.9957
19			1.0000	1.0000	0.9999	0.9997	0.9995	0.9989	0.9980
20						0.9999	0.9998	0.9996	0.9991
21						1.0000	0.9999	0.9998	0.9996
22							1.0000	0.9999	0.9999
23								1.0000	0.9999
24									1.0000

Nature of the Poisson Probability Function



Poisson density functions for different means

the form of the Poisson distribution becomes more and more symmetric, even bell-shaped, as the mean grows large.

Approximation of Binomial Distribution by a Poisson Distribution

- In the case of the binomial, if *n* is quite large and *p is* small, the conditions begin to simulate the continuous space or time implications of the Poisson process

Let X be a binomial random variable with probability distribution $b(x; n, p)$. When $n \to \infty$, $p \to 0$, and $np \overset{n\to\infty}{\longrightarrow} \mu$ remains constant,

$$b(x; n, p) \overset{n\to\infty}{\longrightarrow} p(x; \mu).$$

Example

- Example
  - Suppose the probability that an item will be defective is 0.1
  - Find the probability that a sample of 10 items will contain at most one defective item

Binomial random variable $\binom{10}{0}(.1)^0(.9)^{10} + \binom{10}{1}(.1)^1(.9)^9 = .7361$

Poisson random variable $e^{-1}\frac{1^0}{0!} + e^{-1}\frac{1^1}{1!} = 2e^{-1} \approx .7358$

Wireless Intelligent Networking LAB

Example 5.19: In a certain industrial facility, accidents occur infrequently. It is known that the probability of an accident on any given day is 0.005 and accidents are independent of each other.

(a) What is the probability that in any given period of 400 days there will be an accident on one day?

(b) What is the probability that there are at most three days with an accident?

***Solution*:** Let X be a binomial random variable with $n = 400$ and $p = 0.005$. Thus, $np = 2$. Using the Poisson approximation,

(a) $P(X = 1) = e^{-2}2^1 = 0.271$ and

(b) $P(X \leq 3) = \sum_{x=0}^{3} e^{-2}2^x/x! = 0.857$.

End of Ch. 5

Appendix

- Geometric
  - Proof of mean and variance
  - Proof of Memoryless property
- Poisson
  - Proof of mean and variance

Geometric RV: Mean Proof

- Property
  - Mean $\mu = \dfrac{1}{p}$ $(p>0)$

- Proof

$(q = 1\text{-}p)$

$$E(X) = \sum_{x=0}^{\infty} x f(x) = \sum_{x=0}^{\infty} x p q^{x-1} = p \sum_{x=0}^{\infty} x q^{x-1} \qquad \left(x q^{x-1} = \frac{d}{dq} q^x \right)$$

$$= p \sum_{x=1}^{\infty} \frac{d}{dq} q^x = p \frac{d}{dq} \left(\sum_{x=1}^{\infty} q^x \right) = p \frac{d}{dq} \left(\frac{q}{1-q} \right)$$

$$= \frac{p}{(1-q)^2} = \frac{1}{p}$$

WINL Wireless Intelligent Networking LAB

Geometric RV: Variance Proof

- Property
  - Variance $\sigma^2 = \dfrac{q}{p^2}$ (q = 1-p)

    - Proof $\mathrm{Var(X)} = E[X^2] - E[X]^2$

$$E[X^2] = \sum_{x=1}^{\infty} x^2 f(x) = \sum_{x=1}^{\infty} x^2 p q^{x-1} = \sum_{x=1}^{\infty} x(x+1) p q^{x-1} - \sum_{x=1}^{\infty} x p q^{x-1}$$

$$= p \sum_{x=1}^{\infty} x(x+1) q^{x-1} - \frac{1}{\mathrm{p}} = \mathrm{p}\frac{2}{(1-p)^3} - \frac{1}{p} = \frac{2}{p^2} - \frac{1}{p}$$

$$\boxed{\sum_{n \geq 1} n\,(n+1)\,x^{n-1} = \frac{2}{(1-x)^3}} \quad |x| < 1$$

$$E[X]^2 = \frac{1}{p^2}$$

$$\mathrm{Var(X)} = E[X^2] - E[X]^2 = \frac{2}{p^2} - \frac{1}{p} - \frac{1}{\mathrm{p}^2} = \frac{\mathrm{q}}{\mathrm{p}^2}$$

Memoryless property

- Memoryless
  - Check

$$P(X > n+m \mid X > n) = \frac{P(X > n, X > n+m)}{P(X > n)} = \frac{P(X > n+m)}{P(X > n)}$$

$$P(X > m) = P(X \geq m+1) = \sum_{x=m+1}^{\infty} pq^{x-1} = p \cdot \frac{q^m}{1-q} = q^m$$

$$P(X > n) = q^n, \quad P(X > n+m) = q^{n+m}$$

$$P(X > n+m \mid X > n) = \frac{P(X > n+m)}{P(X > n)} = \frac{q^{n+m}}{q^n} = q^m = P(X > m)$$

Wireless Intelligent Networking LAB

Poisson RV: Mean

- A random variable X, taking on one of the values 0, 1, 2, … , with parameter λ, λ>0 $X \sim \mathrm{Pois}(np)$ $X \sim \mathrm{Pois}(\lambda)$
- The probability mass function

$$p(x\ ;\ \mu) = \frac{e^{-\mu}\mu^{x}}{x\,!},\ x = 0,\ 1,\ 2,\ \ldots$$

- Expectation =μ

Taylor series

$$\sum_{i=0}^{\infty} i \cdot e^{-\mu}\frac{\mu^i}{i!} = \sum_{i=1}^{\infty} i \cdot e^{-\mu}\frac{\mu^i}{i!} = \mu \cdot \sum_{i=1}^{\infty} e^{-\mu}\frac{\mu^{i-1}}{(i-1)!} = \mu \cdot \sum_{j=0}^{\infty} e^{-\mu}\frac{\mu^j}{j!} = \mu$$

Let $j = i$-1

Poisson RV: Variation

- Variation =μ

Let $j = i$-1

$$E[X^2] = \sum_{i=0}^{\infty} i^2 \cdot e^{-\mu} \frac{\mu^i}{i!} = \sum_{j=0}^{\infty} (j+1)^2 \cdot e^{-\mu} \frac{\mu^{j+1}}{(j+1)!} = \sum_{j=0}^{\infty} (j+1)(j+1) \cdot e^{-\mu} \frac{\mu \cdot \mu^j}{(j+1) \cdot j!}$$

$$= \mu \cdot \sum_{j=0}^{\infty} (j+1) \cdot e^{-\mu} \frac{\mu^j}{j!} = \mu \cdot \underbrace{\boxed{\sum_{j=0}^{\infty} j \cdot e^{-\mu} \frac{\mu^j}{j!}}}_{\text{E[X]}} + \mu \cdot \sum_{j=0}^{\infty} \underbrace{\boxed{e^{-\mu} \frac{\mu^j}{j!}}}_{\text{pmf}} = \mu^2 + \mu$$

$$\mathrm{Var}[X] = E[X^2] - E[X]^2 = \mu$$