3.5 Hypergeometric Distribution and Negative Binomial Distribution

○ Hypergeometric and Negative Binomial distributions are both related to Binomial distribution.

Hypergeometric and Negative Binomial Distributions

○ **Binomial distribution** is approximate probability model for sampling without replacement from finite dichotomous (S–F) population provided sample size n is small relative to population size N;

○ **Hypergeometric distribution** is exact probability model for number of S's in sample.

○ **Binomial random variable X** is number of S's when **number n of trials is fixed**, whereas

○ **Negative Binomial distribution** arises from **fixing the number of S's desired** and letting number of trials (n) be random.

Hypergeometric Distribution

Hypergeometric Distribution

Assumptions leading to hypergeometric distribution are as follows:

1. Population or set to be sampled consists of *N* individuals, objects, or elements (a *finite* population).

2. Each individual can be characterized as success (*S*) or failure (*F*), and there are ***M*** **successes** in population.

3. Sample of *n* individuals is selected without replacement in such a way that each subset of size *n* is equally likely to be chosen.

- Random variable of interest is X = number of S's in sample.
- Probability distribution of X depends on the parameters n, M, and N,
- so we wish to obtain $P(X = x) = h(x; n, M, N)$.

Hypergeometric Distribution

$P(X = x) = h(x; n, M, N)$.

Example 3.35

- During a particular period a university's Information Technology office received 20 service orders for problems with printers, of which
  - 8 were laser printers and
  - 12 were inkjet models.
- Sample of 5 of these service orders is to be selected for inclusion in customer satisfaction survey.
- Suppose that 5 are selected in completely random fashion, so that any particular subset of size 5 has the same chance of being selected as does any other subset.
- What is probability that exactly x ($x = 0, 1, 2, 3, 4,$ or 5) of selected service orders were for inkjet printers?

Example 3.35

cont'd

○ Here, population size is $N = 20$, sample size is $n = 5$, and number of S's (inkjet $= S$) and F's in population are $M = 12$ and $N - M = 8$, respectively.

○ Consider value $x = 2$.

○ Because all outcomes (each consisting of 5 particular orders) are equally likely,

sample size (n) population size (N)

$$P(X = 2) = h(2; 5, 12, 20) = \frac{\text{number of outcomes having } X = 2}{\text{number of possible outcomes}}$$

x number of S's (M)

166

Example 3.35

cont'd

○ Number of possible outcomes in experiment is number of ways of selecting 5 from 20 objects without regard to order

$$\text{No. of possible outcomes} = \binom{20}{5}$$

○ To count number of outcomes having $X = 2$,

$$\text{There are } \binom{12}{2} \text{ ways of selecting 2 of inkjet orders}$$

○ for each such way

$$\text{There are } \binom{8}{3} \text{ ways of selecting 3 laser orders to fill out sample}$$

○ From product rule $\text{No. of outcomes with } X = 2 = \binom{12}{2}\binom{8}{3}$

$$P(X = 2) = h(2; 5, 12, 20) = \frac{\text{number of outcomes having } X = 2}{\text{number of possible outcomes}} = \frac{\binom{12}{2}\binom{8}{3}}{\binom{20}{5}} = \frac{77}{323} = 0.238$$

167

Hypergeometric Distribution

In Example 3.35, $n = 5$, $M = 12$, and $N = 20$, so $h(x;\, 5, 12, 20)$ for $x = 0, 1, 2, 3, 4, 5$ can be obtained by substituting these numbers into Equation (3.15).

$$P(X = x) = h(x;\, n, M, N) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

$$P(X = 0) = h(0;\, 5, 12, 20) = \frac{\binom{12}{0}\binom{20-12}{5-0}}{\binom{20}{5}}$$

$$P(X = 1) = h(1;\, 5, 12, 20) = \frac{\binom{12}{1}\binom{20-12}{5-1}}{\binom{20}{5}}$$

$$P(X = 2) = h(2;\, 5, 12, 20) = \frac{\binom{12}{2}\binom{20-12}{5-2}}{\binom{20}{5}}$$

$$P(X = 3) = h(3;\, 5, 12, 20) = \frac{\binom{12}{3}\binom{20-12}{5-3}}{\binom{20}{5}}$$

$$P(X = 4) = h(4;\, 5, 12, 20) = \frac{\binom{12}{4}\binom{20-12}{5-4}}{\binom{20}{5}}$$

$$P(X = 5) = h(5;\, 5, 12, 20) = \frac{\binom{12}{5}\binom{20-12}{5-5}}{\binom{20}{5}}$$

168

Hypergeometric Distribution

- In general, if sample size $\boldsymbol{n}$ is smaller than number of successes in population ($\boldsymbol{M}$), then the largest possible X value is n.

sample size

population size

$$P(X = x) = h(x;\, n, M, N)$$

number of S's

n M

- However, if $M < n$ (e.g., a sample size of 25 and only 15 successes in the population), then $\boldsymbol{X}$ can be at most $\boldsymbol{M}$.

$$\begin{aligned} P(X = x) &= h(x;\, n, M, N) \\ &= h(x;\, 25, 15, N) \end{aligned}$$

169

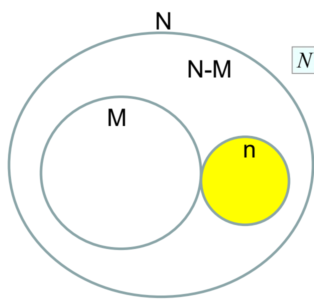
Hypergeometric Distribution

sample size | population size | number of S's

$P(X = x) = h(x; n, M, N)$

- Similarly, whenever number of population failures $(N - M)$ exceeds sample size (n), the smallest possible X value is 0 (since all sampled individuals might then be failures). ⇨ $N - M > n$

Ex. $N=20$, $M=12$, $n=5$ ⇨ $20 - 12 > 5$ ⇨ $8 > 5$



$N - M > n$

smallest possible X value is 0

170

Hypergeometric Distribution

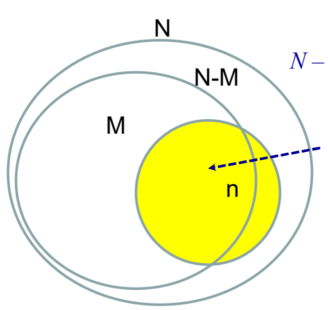
sample size | population size | number of S's

$P(X = x) = h(x; n, M, N)$

- However, if $N - M < n$, the smallest possible X value is $n - (N - M)$.

Ex. $N=20$, $M=18$, $n=5$ ⇨ $20 - 18 < 5$ ⇨ $2 < 5$

smallest possible X value is $n - (N - M) = 5 - (20 - 18) = 3$



$N - M < n$

Smallest possible X value is $n - (N - M)$.

171

Hypergeometric Distribution

sample size

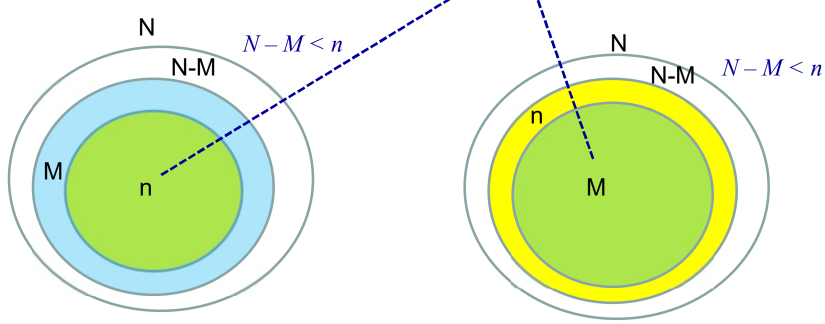
population size

$P(X = x) = h(x; n, M, N)$

number of S's

- Thus, the possible values of X satisfy the restriction

$$\max(0, n - (N - M)) \le x \le \min(n, M).$$



- An argument parallel to that of previous example gives pmf of X.

172

Hypergeometric Distribution

Proposition

If X is number of S's in a completely random sample of size n drawn from population consisting of M S's and $(N - M)$ F's, then probability distribution of X, called **Hypergeometric Distribution**, is given by

Number of Outcomes having X=x

$$P(X = x) = h(x; n, M, N) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} \qquad \textbf{(3.15)}$$

Number of possible outcomes

for x, an integer, satisfying $\max(0, n - N + M) \le x \le \min(n, M)$.

173

Example 3.36

- Five individuals from animal population thought to be near extinction in a certain region have been caught, tagged, and released to mix into population
- After they have had an opportunity to mix, random sample of 10 of these animals is selected.
- Let X = the number of tagged animals in the second sample
- If there are actually 25 animals of this type in region,
- what is probability that
  a) X=2?
  b) X<= 2?

174

Example 3.36

a) P(X=2)

175

Example 3.36

b) P(X<= 2)

Hypergeometric Distribution

As in binomial case, there are simple expressions for $E(X)$ and $V(X)$ for hypergeometric random variable's.

Proposition

Mean and Variance of Hypergeometric rv X having pmf $h(x; n, M, N)$ are

$$E(X) = n \cdot \frac{M}{N} \qquad V(X) = \left(\frac{N-n}{N-1}\right) \cdot n \cdot \frac{M}{N} \cdot \left(1 - \frac{M}{N}\right)$$

Ratio M/N is the proportion of S's in population.

If we replace M/N by p in $E(X)$ and $V(X)$, we get

$$E(X) = np \qquad V(X) = \left(\frac{N-n}{N-1}\right) \cdot np \cdot (1-p) \qquad \textbf{(3.16)}$$

Example 3.37 (from example 3.36)

○ Five individuals from an animal population thought to be near extinction in a certain region have been caught, tagged, and released to mix into the population.

○ After they have had an opportunity to mix, a random sample of 10 of these animals is selected.

○ Let X = the number of tagged animals in second sample.

○If there are actually 25 animals of this type in the region,

what is the ***E*(*X*)** and ***V*(*X*)**?

185

Example 3.37

cont'd

In the animal-tagging example,

$n = 10$, $M = 5$, and $N = 25$, so $\frac{M}{N} = p = \frac{5}{25} = 0.2$

$\boxed{E(X) = n \cdot \frac{M}{N}}$ ⇨ $\boxed{E(X) = np}$ ⇨ $\boxed{E(X) = 10(0.2) = 2}$

$\boxed{V(X) = \left(\frac{N-n}{N-1}\right) \cdot n \cdot \frac{M}{N} \cdot \left(1 - \frac{M}{N}\right)}$ ⇨ $\boxed{V(X) = \left(\frac{N-n}{N-1}\right) \cdot np \cdot (1-p)}$

$$\begin{aligned} V(X) &= \left(\frac{25-10}{25-1}\right) \cdot 10(0.2) \cdot (1-0.2) \\ &= \left(\frac{15}{24}\right) \cdot 10(0.2) \cdot (0.8) \\ &= (0.625)(1.6) \\ &= 1 \end{aligned}$$

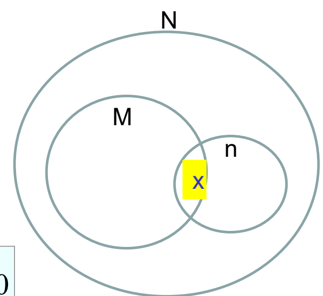
186

Example 3.37

cont'd

- Suppose population size N is not actually known, so value x is observed and we wish to estimate N.
- It is reasonable to equate the observed sample proportion of S's, x/n, with population proportion, M/N, giving the estimate

$$\frac{x}{n} = \frac{M}{N} \Rightarrow \hat{N} = \frac{M \cdot n}{x}$$



- If $M = 100$, $n = 40$, and $x = 16$, then

$$\hat{N} = \frac{M \cdot n}{x} = \frac{(100)(40)}{16} = 250$$

188

Negative Binomial Distribution

190

Negative Binomial Distribution

- Negative binomial random variable and distribution are based on experiment satisfying the following conditions:

1. Experiment consists of a sequence of independent trials.

2. Each trial can result in either a success (S) or a failure (F).

3. Probability of success is constant from trial to trial, so for $i = 1, 2, 3, \ldots$

4. Experiment continues (trials are performed) until a total of r successes have been observed, where r is a specified positive integer.

191

Negative Binomial Distribution

- Random variable of interest is

จำนวนการทดลองที่ไม่สำเร็จก่อนหน้าความสำเร็จครั้งที่ r

$\boxed{X = \text{the number of failures that precede the } r^{th} \text{ success}}$

- X is called a **Negative Binomial Random Variable** because, in contrast to binomial random variable, number of successes is fixed and the number of trials is random.

Binomial Random Variable กำหนดจำนวนครั้งของการทดลอง (Trial) ที่แน่นอน
จำนวนครั้งของความสำเร็จขึ้นอยู่กับความสนใจ

192

Negative Binomial Distribution

จำนวนครั้งของความสำเร็จ

- Possible values of X are 0, 1, 2, ...
- Let $nb(x; r, p)$ denote pmf of X.
- Consider $nb(7; 3, p) = P(X = 7)$,

  probability that exactly 7 F's occur before the 3rd S.

- In order for this to happen, the 10th trial must be an S and there must be exactly 2 S's among the first 9 trials. Thus

$$nb(7; 3, p) = \left\{\binom{9}{2} \cdot p^2(1-p)^7\right\} \cdot p = \binom{9}{2} \cdot p^3(1-p)^7 \Rightarrow \binom{\overbrace{7+3-1}^{x+r-1}}{\underbrace{3-1}_{r-1}} \cdot \overbrace{p^3}^{r}(1-p)^{\overbrace{7}^{x}}$$

- Generalizing this line of reasoning gives the following formula for negative binomial pmf.

Negative Binomial Distribution

Proposition

The pmf of negative binomial random variable X with parameters r = number of S's and $p = P(S)$ is

$$nb(x; r, p) = \binom{x+r-1}{r-1} p^r(1-p)^x \qquad x = 0, 1, 2, \ldots$$

จำนวนครั้งของความไม่สำเร็จ

จำนวนครั้งของความสำเร็จ

ความน่าจะเป็นของความสำเร็จ

Negative Binomial Distribution

Example 3.38

- A pediatrician wishes to recruit 5 couples, each of whom is expecting their first child, to participate in a new natural childbirth regimen.
- Let $p = P$(a randomly selected couple agrees to participate).
- If $p = 0.2$, what is **probability** that **15 couples** must be asked **before 5 are found** who agree to participate?
- That is, with $S = \{\text{agrees to participate}\}$,
- what is **probability** that **10 *F*'s occur before the fifth *S***?
- Substituting $r = 5$, $p = 0.2$, and $x = 10$ into $nb(x; r, p)$ gives

$$nb(x; r, p) = \binom{x+r-1}{r-1} p^r (1-p)^x \quad x = 0, 1, 2, \ldots$$

$$nb(10; 5, 0.2) = \binom{14}{4} (0.2)^5 (1-0.2)^{10} = 0.034$$

Example 3.38

Solution

Example 3.38

$nb(x; r, p) = \binom{x+r-1}{r-1} p^r (1-p)^x \quad x = 0, 1, 2, \ldots$

cont'd

○ Probability that **at most** 10 F's are observed (at most 15 couples are asked) is

$$nb(x; r, p) = \binom{x+r-1}{r-1} p^r (1-p)^x \quad x = 0, 1, 2, \ldots$$

Negative Binomial Distribution

- In the special case $r = 1$, the pmf is

$$nb(x; 1, p) = p(1-p)^x \quad x = 0, 1, 2, \ldots \qquad \textbf{(3.18)}$$

- In Example 3.12 (slide 39), we derived pmf for the number of trials necessary to obtain the first S, and pmf there is similar to Expression (3.18).

- Both $X =$ number of F's and $Y =$ number of trials ($= 1 + x$) are referred to in the literature as **Geometric Random Variables,** and pmf in Expression (3.18) is called **Geometric Distribution.**

202

Negative Binomial Distribution

Proposition

If $\boldsymbol{X}$ is a negative binomial random variable with pmf $nb(x; r, p)$, then

$$E(X) = \frac{r(1-p)}{p} \qquad V(X) = \frac{r(1-p)}{p^2}$$

204

3.6 Poisson Probability Distribution

217

Poisson Probability Distribution

- Binomial, Hypergeometric, and Negative Binomial Distributions were all derived by starting with experiment consisting of trials or draws and applying laws of probability to various outcomes of experiment.

- There is no simple experiment on which Poisson distribution is based, though we will shortly describe how it can be obtained by certain limiting operations.

- In contrast to Binomial and Hypergeometric Distributions, Poisson Distribution spreads probability over *all* non-negative integers, an infinite number of possibilities.

218

Poisson Probability Distribution

Definition

Discrete random variable ***X*** is said to have **Poisson Distribution** with parameter λ ($\lambda > 0$) if **pmf of *X*** is

e represents base of natural logarithm system;
its numerical value is approximately 2.71828.

$$p(x;\lambda) = \frac{e^{-\lambda}\cdot\lambda^x}{x!} \qquad x = 0, 1, 2, 3, \ldots$$

- Value of λ is frequently rate per unit time or per unit area
- Because λ must be positive; $p(x;\lambda) > 0$ for all possible x values

219

Poisson Probability Distribution

- The fact that $\Sigma\, p(x;\lambda) = 1$ is consequence of the Maclaurin series expansion of e^{λ} , which appears in most calculus texts :

$$e^{\lambda} = 1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \cdots = \sum_{x=0}^{\infty}\frac{\lambda^x}{x!} \qquad (3.19)$$

- If the two extreme terms in (3.19) are multiplied by $e^{-\lambda}$ and then this quantity is moved inside the summation on the far right, the result is

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

- which shows that $p(x;\lambda)$ fulfills the second condition necessary for specifying pmf

220

Example 3.39 $p(x;\lambda) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}$ $x = 0, 1, 2, 3, \ldots$

- Let X denote the number of creatures of a particular type captured in a trap during a given time period.
- Suppose that X has Poisson distribution with $\lambda = 4.5$, so on average traps will contain 4.5 creatures.
- Probability that a trap contains **exactly** **five** creatures is
- Probability that a trap has **at most** **five** creatures is

Poisson Distribution as a Limit

Poisson Distribution as a Limit

- The rationale for using Poisson distribution in many situations is provided by the following proposition.

Proposition

- Suppose that in Binomial pmf $b(x; n, p)$,
- we let $n \to \infty$ and $p \to 0$ in such a way that np approaches a value $\lambda > 0$.
- Then $b(x; n, p) \to p(x; \lambda)$.

- According to this proposition, *in any binomial experiment in which **n is large** and **p is small**,* $b(x; n, p) \approx p(x; \lambda)$,
- *where* $\lambda = np$.
- As a rule of thumb, this approximation can safely be applied

if $n > 50$ **and** $np < 5$.

224

Example 3.40

- If publisher of nontechnical books takes great pains to ensure that its books are free of typographical errors,
- so that probability of any given page containing **at least one** such **error** is 0.005 and
- errors are independent from page to page,
- what is probability that one of its 400-page novels will contain
  1) **exactly** one page with errors?
  2) **at most** three pages with errors?

- With S denoting page containing at least one error and F an error-free page,
- the number X of pages containing at least one error is a binomial random variable with $n = 400$ and $p = 0.005$, so

$$\boxed{\lambda = np = (400)(0.005) = 2}$$

225

Example 3.40

$p(x; \lambda) = \frac{e^{-\lambda} \cdot \lambda^x}{x!} \quad x = 0, 1, 2, 3, \ldots$

cont'd

What is probability that one of its 400-page novels will contain **exactly** one page with errors?

Binomial value is $b(1; 400, 0.005) = 0.270669$,

so approximation is very good.

Example 3.40

$p(x; \lambda) = \frac{e^{-\lambda} \cdot \lambda^x}{x!} \quad x = 0, 1, 2, 3, \ldots$

cont'd

What is probability that one of its 400-page novels will contain **at most** three pages with errors?

and this again is quite close to the Binomial value $P(X \leq 3) = 0.8576$

Poisson Distribution as a Limit

Table 3.2 shows Poisson distribution for $\lambda = 3$ along with three binomial distributions with $np = 3$

x	$n = 30, p = .1$	$n = 100, p = .03$	$n = 300, p = .01$	Poisson, $\lambda = 3$
0	0.042391	0.047553	0.049041	0.049787
1	0.141304	0.147070	0.148609	0.149361
2	0.227656	0.225153	0.224414	0.224042
3	0.236088	0.227474	0.225170	0.224042
4	0.177066	0.170606	0.168877	0.168031
5	0.102305	0.101308	0.100985	0.100819
6	0.047363	0.049610	0.050153	0.050409
7	0.018043	0.020604	0.021277	0.021604
8	0.005764	0.007408	0.007871	0.008102
9	0.001565	0.002342	0.002580	0.002701
10	0.000365	0.000659	0.000758	0.000810

Table 3.2 Comparing the Poisson and Three Binomial Distributions

Poisson Distribution as a Limit

- Figure 3.8 (from S-Plus) plots Poisson along with first two Binomial Distributions.
  - Approximation is of limited use for $n = 30$, the accuracy is better for $n = 100$ and much better for $n = 300$.

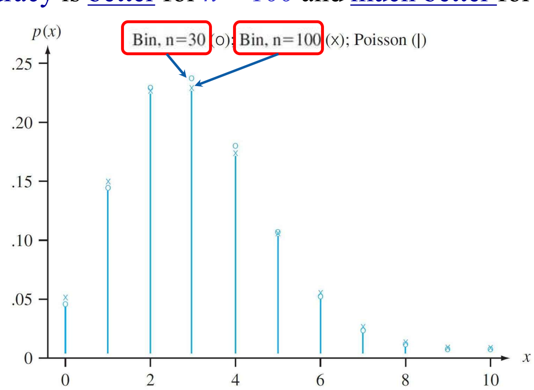


Figure 3.8 Comparing a Poisson and two binomial distributions

Poisson Distribution as a Limit

Appendix Table A.2 exhibits cdf $F(x; \lambda)$ for $\lambda = 0.1, 0.2, \ldots, 1, 2, \ldots, 10, 15$, and 20.

Table A.2 Cumulative Poisson Probabilities

$$F(x; \lambda) = \sum_{y=0}^{x} \frac{e^{-\lambda}\lambda^y}{y!}$$

		λ									
		.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
	0	.905	.819	.741	.670	.607	.549	.497	.449	.407	.368
	1	.995	.982	.963	.938	.910	.878	.844	.809	.772	.736
	2	1.000	.999	.996	.992	.986	.977	.966	.953	.937	.920
x	3		1.000	1.000	.999	.998	.997	.994	.991	.987	.981
	4				1.000	1.000	1.000	.999	.999	.998	.996
	5							1.000	1.000	1.000	.999
	6										1.000

Poisson Distribution as a Limit

Table A.2 Cumulative Poisson Probabilities (cont.)

$$F(x; \lambda) = \sum_{y=0}^{x} \frac{e^{-\lambda}\lambda^y}{y!}$$

		λ										
		2.0	3.0	4.0	5.0	6.0	7.0	8.0	9.0	10.0	15.0	20.0
	0	.135	.050	.018	.007	.002	.001	.000	.000	.000	.000	.000
	1	.406	.199	.092	.040	.017	.007	.003	.001	.000	.000	.000
	2	.677	.423	.238	.125	.062	.030	.014	.006	.003	.000	.000
	3	.857	.647	.433	.265	.151	.082	.042	.021	.010	.000	.000
	4	.947	.815	.629	.440	.285	.173	.100	.055	.029	.001	.000
	5	.983	.916	.785	.616	.446	.301	.191	.116	.067	.003	.000
	6	.995	.966	.889	.762	.606	.450	.313	.207	.130	.008	.000
	7	.999	.988	.949	.867	.744	.599	.453	.324	.220	.018	.001
	8	1.000	.996	.979	.932	.847	.729	.593	.456	.333	.037	.002
	9		.999	.992	.968	.916	.830	.717	.587	.458	.070	.005
	10		1.000	.997	.986	.957	.901	.816	.706	.583	.118	.011
	11			.999	.995	.980	.947	.888	.803	.697	.185	.021
	12			1.000	.998	.991	.973	.936	.876	.792	.268	.039
	13				.999	.996	.987	.966	.926	.864	.363	.066
	14				1.000	.999	.994	.983	.959	.917	.466	.105
	15					.999	.998	.992	.978	.951	.568	.157
	16					1.000	.999	.996	.989	.973	.664	.221
	17						1.000	.998	.995	.986	.749	.297
	18							.999	.998	.993	.819	.381
x	19							1.000	.999	.997	.875	.470
	20								1.000	.998	.917	.559
	21									.999	.947	.644
	22									1.000	.967	.721
	23										.981	.787
	24										.989	.843
	25										.994	.888
	26										.997	.922
	27										.998	.948
	28										.999	.966
	29										1.000	.978
	30											.987
	31											.992
	32											.995
	33											.997
	34											.999
	35											.999
	36											1.000

Poisson Distribution as a Limit

For example, if $\lambda = 2$, then $P(X \leq 3) = F(3; 2) = 0.857$ as in example 3.40, whereas $P(X = 3) = F(3; 2) - F(2; 2) = 0.857 - 0.677 = 0.180$.

Table A.2 Cumulative Poisson Probabilities *(cont.)*

$$F(x; \lambda) = \sum_{y=0}^{x} \frac{e^{-\lambda}\lambda^y}{y!}$$

		λ										
		2.0	3.0	4.0	5.0	6.0	7.0	8.0	9.0	10.0	15.0	20.0
	0	.135	.050	.018	.007	.002	.001	.000	.000	.000	.000	.000
	1	.406	.199	.092	.040	.017	.007	.003	.001	.000	.000	.000
	2	.677	.423	.238	.125	.062	.030	.014	.006	.003	.000	.000
x	3	.857	.647	.433	.265	.151	.082	.042	.021	.010	.000	.000
	4	.947	.815	.629	.440	.285	.173	.100	.055	.029	.001	.000
	5	.983	.916	.785	.616	.446	.301	.191	.116	.067	.003	.000
	6	.995	.966	.889	.762	.606	.450	[illegible]	[illegible]	[illegible]		

Alternatively, many statistical computer packages will generate $p(x; \lambda)$ and $F(x; \lambda)$ upon request.

Mean and Variance of X

Mean and Variance of X

- Since $b(x; n, p) \rightarrow p(x; \lambda)$ as $n \infty$, $p \rightarrow 0$, $np \rightarrow \lambda$,
- Mean and Variance of binomial variable should approach those of Poisson variable.
- These limits are $np \rightarrow \lambda$ and $np(1-p) \rightarrow \lambda$.

Proposition

If X has a Poisson distribution with parameter λ, then

$$E(X) = V(X) = \lambda$$

These results can also be derived directly from the definitions of mean and variance.

236

Example 3.41

Example 3.39 continued…

Both **expected** number of creatures trapped and **variance** of the number trapped equal **4.5**, and

$$E(X) = V(X) = \lambda$$

Standard deviation (SD)

$$\begin{aligned} \sigma_X &= \sqrt{\lambda} \\ &= \sqrt{4.5} \\ &= 2.12 \end{aligned}$$

237

Poisson Process

Poisson Process

- A very important application of Poisson distribution arises in connection with **occurrence of events** of some type **over time**.
- Events of interest might be
  - visits to a particular website,
  - pulses of some sort recorded by counter,
  - email messages sent to a particular address,
  - accidents in an industrial facility, or
  - cosmic ray showers observed by astronomers at a particular observatory.

จำนวนลูกค้าที่มาถึงยัง Counter ให้บริการ
จำนวนของ Call ที่มาถึงยัง Telephone Exchange
จำนวนของ Packet ที่มาถึงยัง Queue

Poisson Process

We make following assumptions about the way in which events of interest occur:

1. There exists parameter $\alpha > 0$ such that for any short time interval of length Δt,
   probability that exactly one event occurs is received is $\alpha \cdot \Delta t + o(\Delta t)$*

2. Probability of more than one event occurring during Δt is $o(\Delta t)$
   [which, along with Assumption 1, implies that probability of no events during Δt is $1 - \alpha \cdot \Delta t - o(\Delta t)$]

3. The number of events occurring during the time interval Δt is independent of the number that occur prior to this time interval.

- * Quantity is $o(\Delta t)$ (read "little o of delta t" is as Δt *approaches* 0, so does $o(\Delta t)/\Delta t$
- That is, $o(\Delta t)$ is even more negligible (approaches 0 faster) than Δt itself.
- The $(\Delta t)^2$ has this property

Poisson Process

- Informally, Assumption 1 says that for a short interval of time, probability of receiving single event (occurring) is approximately proportional to the length of time interval,

  where α is the constant of proportionality.

- Now let $P_k(t)$ denote probability that k events will be observed during any particular time interval of length t.

Poisson Process

Proposition

$$P_k(t) = e^{-\alpha t} \cdot \frac{(\alpha t)^k}{k!}$$

- so that the number of events during time interval of length t is Poisson random variable with parameter $\lambda = \alpha t$.
- Expected number of events during any such time interval is then αt, so
- expected number during a unit interval of time is α.

- **Occurrence of events over time** as described is called a ***Poisson Process***; **parameter α** specifies ***rate*** for process.

242

Example 42

- Suppose pulses arrive at counter at average rate of **six per minute**.
- Find probability that in 0.5-min interval **at least** one pulse is received.

Solution

243

Example

- Customers arrive to a bank according to Poisson Process having a constant average rate of 8.6 customers per hour.
- Suppose we begin observing the bank at some point in time.

a) What is the expected value of the number of customers that arrive in the first 30 min.?

a) What is the probability that 3 customers arrive in the first 30 min.?

245

End of Chapter 3

246