ENGG 319

Probability & Statistics for Engineers

Section #05

Discrete
Probability Distribution

L01

Dr. Sameh Nassar

F16

Discrete Probability Distributions

- Discrete probability distributions can be represented graphically by a histogram (*Section #03*), in a tabular form, or by a formula.
- Usually, observations from different experiments will have the same type of behavior, and hence the corresponding discrete random variables can be described by the same probability distribution.

Some Discrete Probability Distributions:

- Uniform
- Binomial
- Multinomial
- Hypergeometric
- Poisson

Discrete Uniform Distribution

- If the random variable *X* assumes the values $x_1, x_2, \ldots, x_k$, with equal probabilities, then the discrete uniform distribution is given by:

$$f(x;k) = \frac{1}{k} \quad , \qquad x = x_1, x_2, \ldots, x_k$$

- The mean and variance of the discrete uniform distribution *f(x; k)* are:

$$\mu_x = \frac{1}{k}\sum_{i=1}^{k} x_i \qquad \sigma_x^2 = \frac{1}{k}\sum_{i=1}^{k}(x_i - \mu_x)^2$$

Example #1

- If a fair die is tossed once, obtain the following:

(a) The probability distribution.

(b) The mean.

(c) The variance.

(d) The corresponding histogram.

Example #1 (Sol.)

$S = \{ 1, 2, 3, 4, 5, 6 \}$ ➡ $k = 6$

(a) $$f(x;k) = \frac{1}{k}$$

➡ $$f(x;k) = \frac{1}{6} \quad , \qquad x = 1,2,3,4,5,6$$

(b) $$\mu_x = \frac{1}{k}\sum_{i=1}^{k} x_i$$ ➡ $$\mu_x = \frac{1}{6}\sum_{i=1}^{6} x_i$$

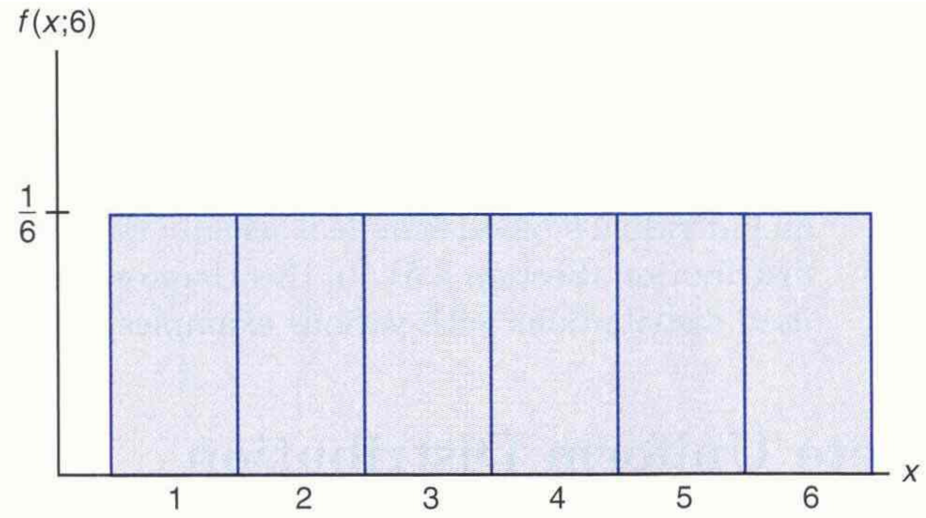
$$\mu_x = \frac{1}{6}(1+2+3+4+5+6) = \frac{7}{2} = 3.5$$

Example #1 (Sol.)

(c) $$\sigma_x^2 = \frac{1}{k}\sum_{i=1}^{k}(x_i - \mu_x)^2 \Rightarrow \sigma_x^2 = \frac{1}{6}\sum_{i=1}^{6}(x_i - 3.5)^2$$

$$\sigma_x^2 = \frac{1}{6}\left[(1-3.5)^2 + (2-3.5)^2 (3-3.5)^2 (4-3.5)^2 (5-3.5)^2 (6-3.5)^2\right]$$
$$= 2.92$$

(d)



Binomial Distribution

- The probability distribution of the binomial random variable *X* (that denotes the number of successes in n ***independent trials***), with the probability of success ***p*** and ***q=1-p*** is the probability of failure is:

$$P(X = x) \quad = f(x) \quad = b(x; n, p) = \binom{n}{x} . p^{x} . q^{n-x}, \quad x = 0, 1, 2, \cdots, n$$

- Thus, for each trial we have only 2 possible outcomes.

- The mean and variance of the binomial distribution *b(x;n,p)* are:

$$\mu_x = np \qquad \sigma_x^2 = npq$$

Example #2

Using data from Section #03 (Example #2)

- If a coin is flipped ***three*** times, and given that the probability of a head outcome is 50%, determine the following:

(a) The probability distribution of the random variable X representing the number of heads.

(b) The probability of the occurrence of 3 heads.

(c) The probability of the occurrence of 2 heads.

(d) The mean and standard deviation of X.

from Section #03 (Example #2 solution)

x	$f(x)$
0	1/8
1	3/8
2	3/8
3	1/8

$$P(X=3)=f(3)=\frac{1}{8}$$

Example #2 (Sol.)

(a) $$P(X = x) = f(x) = b(x; n, p) = \binom{n}{x} . p^x . q^{n-x}, \quad x = 0,1,2,3$$

(b) n = 3 $\quad p = 0.50 \quad q = 1 - 0.50 = 0.50$

$$P(X = 3) = f(3) = b\left(3; 3, \frac{1}{2}\right) = \binom{3}{3} . \left(\frac{1}{2}\right)^3 . \left(\frac{1}{2}\right)^{3-3}$$

$$= \frac{3!}{3!.0!} . \left(\frac{1}{2}\right)^3 . \left(\frac{1}{2}\right)^0 = \frac{1}{8} = 0.125$$

Example #2 (Sol.)

(c) $n = 3$ $p = 0.50$ $q = 1 - 0.50 = 0.50$

$$P(X=2) = f(2) = b\left(2;3,\frac{1}{2}\right) = \binom{3}{2}\cdot\left(\frac{1}{2}\right)^{2}\cdot\left(\frac{1}{2}\right)^{3-2} = \frac{3!}{2!.1!}\cdot\left(\frac{1}{2}\right)^{2}\cdot\left(\frac{1}{2}\right)^{1} = \frac{3}{1}*\frac{1}{8} = \frac{3}{8}$$

(d) $\mu_x = np$ $\sigma_x^2 = npq$

$$\mu_x = 3*\frac{1}{2} = \frac{3}{2}$$

$$\sigma_x^2 = 3*\frac{1}{2}*\frac{1}{2} = \frac{3}{4}$$

$$\sigma_x = \sqrt{\frac{3}{4}} = 0.866$$

Example #3

- Using the information of Example #2, what is the probability of obtaining at least 2 heads?

n = 3

$$p = 0.50$$

$$q = 1 - 0.50 = 0.50$$

$$P(X \geq 2) = ?$$

$$P(X \geq 2) = 1 - P(X < 2) = 1 - \sum_{x=0}^{1} b\left(x;3,\frac{1}{2}\right)$$

$$\sum_{x=0}^{1} b\left(x;3,\frac{1}{2}\right) = \frac{3!}{0!.3!}.\left(\frac{1}{2}\right)^0.\left(\frac{1}{2}\right)^3 + \frac{3!}{1!.2!}.\left(\frac{1}{2}\right)^1.\left(\frac{1}{2}\right)^2$$

$$= 1*1*\frac{1}{8} + 3*\frac{1}{2}*\frac{1}{4} \quad = \quad \frac{4}{8} = \frac{1}{2}$$

$$P(X \geq 2) = 1 - \frac{1}{2} = \frac{1}{2}$$

Binomial Sums (Table A.1)

- Frequently, we are interested in problems involving the computation of $P(X < r)$ or $P(a \leq X \leq b)$.
- For this purpose, we have some of the binomial sums available (Table A.1 in your textbook), where:

$$B(r; n, p) = \sum_{x=0}^{r} b(x; n, p)$$

- For Table A.1, the sums are available for a range of n =1-20, and some selected values of p (0.1 – 0.9) and some corresponding values of r.

Example #4

- Using the information of Example #2, what is the probability of obtaining at least 2 heads using Table A.1?

n = 3 $p = 0.50$ $P(X \geq 2) = ?$

$$P(X \geq 2) = 1 - P(X < 2) = 1 - \sum_{x=0}^{1} b(x;3,0.5)$$

$\boldsymbol{x}$ (or $\boldsymbol{r}$) = 1

n=3

	p																		
r	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95
0	0.857	0.729	0.614	0.512	0.422	0.343	0.275	0.216	0.166	0.125	0.091	0.064	0.043	0.027	0.016	0.008	0.003	0.001	0.000
1	0.993	0.972	0.939	0.896	0.844	0.784	0.718	0.648	0.575	0.500	0.425	0.352	0.282	0.216	0.156	0.104	0.061	0.028	0.007
2	1.000	0.999	0.997	0.992	0.984	0.973	0.957	0.936	0.909	0.875	0.834	0.784	0.725	0.657	0.578	0.488	0.386	0.271	0.143
3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

$$P(X \geq 2) = 1 - 0.500 = 0.50$$

Example #5

- Using the information of Example #3, what is the probability of obtaining 2 heads using Table A.1?

n = 3 $p = 0.50$ $P(X = 2) = ?$

$$P(X = 2) = \sum_{x=0}^{2} b(x;3,0.5) - \sum_{x=0}^{1} b(x;3,0.5)$$

n=3

r	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95
0	0.857	0.729	0.614	0.512	0.422	0.343	0.275	0.216	0.166	0.125	0.091	0.064	0.043	0.027	0.016	0.008	0.003	0.001	0.000
1	0.993	0.972	0.939	0.896	0.844	0.784	0.718	0.648	0.575	0.500	0.425	0.352	0.282	0.216	0.156	0.104	0.061	0.028	0.007
2	1.000	0.999	0.997	0.992	0.984	0.973	0.957	0.936	0.909	0.875	0.834	0.784	0.725	0.657	0.578	0.488	0.386	0.271	0.143
3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

$$P(X = 2) = 0.875 - 0.500 = 0.375 = 3/8$$

Multinomial Distribution

- If a given trial can result in k outcomes with probabilities p_1, p_2, ,..., p_k, then for n independent trials:

$$f(x_1, x_2 \cdots x_k; p_1, p_2, \cdots p_k, n) =$$

$$\binom{n}{x_1, x_2, \cdots, x_k} . p_1^{x_1} . p_2^{x_2} \cdots p_3^{x_3}$$

$$= \left(\frac{n!}{x_1! x_2! \cdots x_k!} \right) p_1^{x_1} . p_2^{x_2} \cdots p_3^{x_3}$$

$$\sum_{i=1}^{k} x_i = n$$

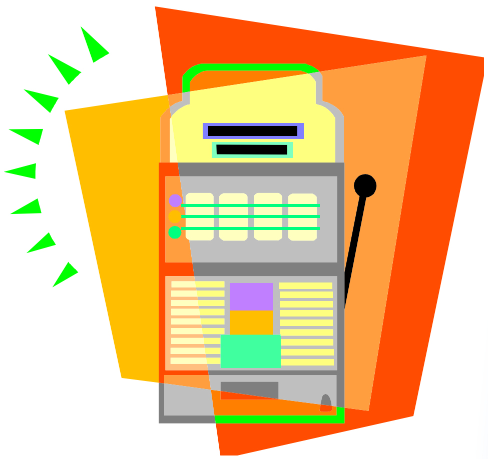
$$\sum_{i=1}^{k} p_i = 1$$

Example #6

- Assume the probability of a component produced by a certain machine being too small is 10%, too big is 15% and within the accepted dimensions is 75%. If seven components produced by that machine are selected randomly, what is the probability of the following.

(a) having four accepted, one too small and two too big components.

(b) having four accepted components.



Example #6 (Sol.)

(a) $$f\left(x_1, x_2 \cdots x_k; p_1, p_2, \cdots p_k, n\right) = \left(\frac{n!}{x_1! x_2! \cdots x_k!}\right) p_1^{x_1} . p_2^{x_2} \cdots p_3^{x_3}$$

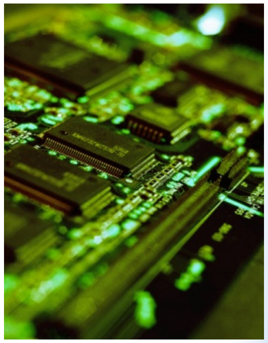
$n_1 = 4$ $n_2 = 1$ $n_3 = 2$ $n = 7$

$p_1 = 0.75$ $p_2 = 0.10$ $p_3 = 0.15$

$$f\left(x_1, x_2, x_3; p_1, p_2, p_3, n\right) = \left(4, \ 1, \ 2 \ ; \ 0.75, \ 0.10, \ 0.15, \ 7\right)$$

$$= \left(\frac{7!}{4!.1!.2!}\right)(0.75)^4.(0.10)^1.(0.15)^2$$

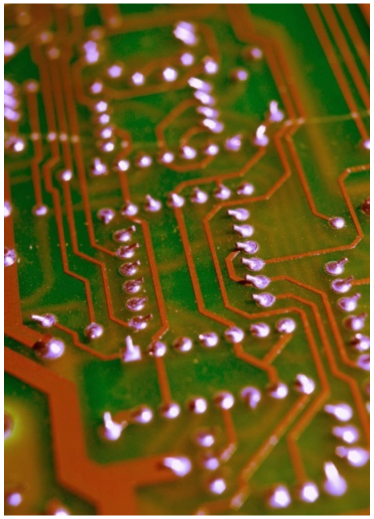
$$= \frac{7*6*5}{1*2*1}(0.75)^4.(0.10)^1.(0.15)^2 = 0.0747 \cong 0.08$$



Example #6 (Sol.)

(b)

$$P(X=x) \quad = b(x;n,p) = \binom{n}{x}.p^{x}.q^{n-x}, \quad x=4$$



n = 7 $\quad p = 0.75 \quad q = 1 - 0.75 = 0.25$

$$P(X=4) = b(4;7,0.75) = \binom{7}{4}.(0.75)^4.(0.25)^{7-4}$$

$$P(X=4) = \left(\frac{7!}{4!.3!}\right).(0.75)^4.(0.25)^3 = \frac{7*6*5}{3*2*1}.(0.75)^4.(0.25)^3$$

$$\cong 0.17$$

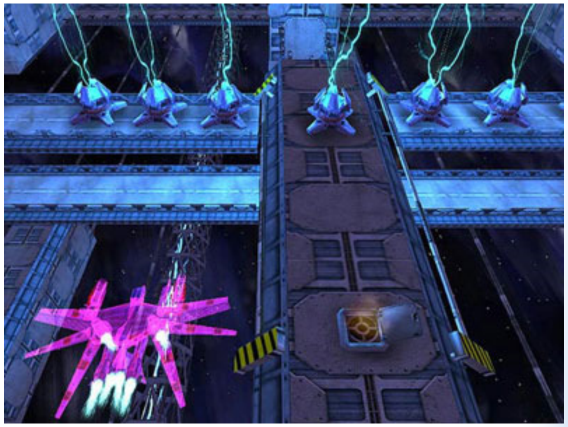
HW: Solve (b) using Table A.1!

Hypergeometric Distribution

- The probability distribution of the ***hypergeometric*** random variable *X*, the number of successes in a random sample of size ***n*** selected from ***N*** items of which ***k*** are labeled success and ***N-k*** labeled failure, is:

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

$$\max\{0,n-(N-k)\}\le x\le \min\{n,k\}$$

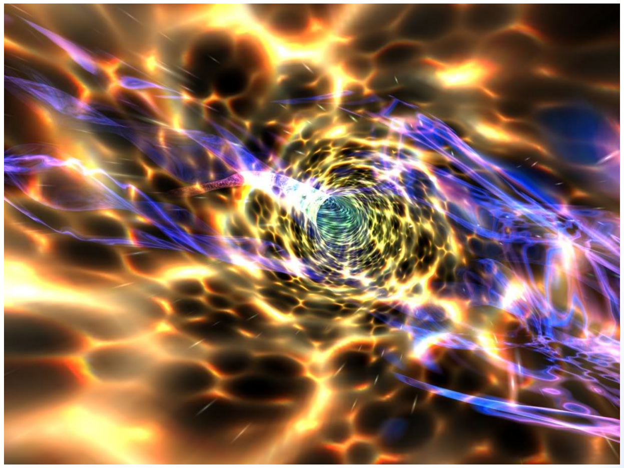


Hypergeometric Distribution

- The mean and variance of the ***hypergeometric*** distribution *h(x; N, n, k)* are:

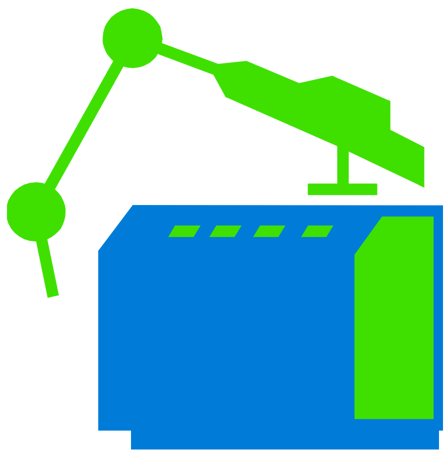
$$\mu_x = \frac{nk}{N}$$

$$\sigma_x^2 = \frac{N-n}{N-1} \cdot n \cdot \frac{k}{N}\left(1 - \frac{k}{N}\right)$$



Example #7

- A certain machine is producing 10 components in an hour. If it is known that from these 10 components, eight are within the accepted dimensions.
- If three components produced by that machines are selected randomly in a one of the production hours, what is the probability of having two accepted and one rejected?



Example #7 (Sol.)

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

N = 10

k = 8

x = 2

n = 3

$$h(2;10,3,8)=\frac{\binom{8}{2}\binom{2}{1}}{\binom{10}{3}}=\frac{\frac{8!}{2!.6!}*\frac{2!}{1!.1!}}{\frac{10!}{3!.7!}}=\frac{\frac{8*7}{2*1}*\frac{2}{1}}{\frac{10*9*8}{3*2*1}}$$

≈ 0.47

Example #8

- A certain machine is producing 1000 components to be packed in a single box. It is known that 80% of the components in each box are within the accepted dimensions.
- If three components are selected randomly from the box, what is the probability of having two components within the accepted dimensions and one that is not?

(a) Use the hypergeometric distribution first.

(b) Then, use the binomial distribution.

(c) Compare the results of (a) and (b) and then comment on the results.

Example #8 (Sol.)

(a)

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

N = 1000

k = 0.8*1000 = 800

x = 2

n = 3

$$h(2;1000,3,800)=\frac{\binom{800}{2}\binom{200}{1}}{\binom{1000}{3}}=\frac{\frac{800!}{2!.798!}*\frac{200!}{1!.199!}}{\frac{1000!}{3!.997!}}=\frac{\frac{800*799}{2*1}*\frac{200}{1}}{\frac{1000*999*998}{3*2*1}}$$

= 0.3847

Example #8 (Sol.)

(b)

$$P(X = x) = b(x; n, p) = \binom{n}{x}.p^{x}.q^{n-x}, \quad x = 2$$

n = 3 $\quad p = 0.80 \quad q = 1 - 0.80 = 0.20$

$$P(X = 2) = b(2;3,0.80) = \binom{3}{2}.(0.8)^{2}.(0.2)^{3-2}$$

$$= \frac{3!}{2!.1!}.(0.8)^{2}.(0.2)^{1} = 0.384$$

(c) $h(2;1000,3,800) \approx b(2;3,0.80)$

This is due to the very small sample size ***n*** compared to the large population ***N*** (i.e. ***N*** >>>> ***n***)

Multivariate Hypergeometric Distribution

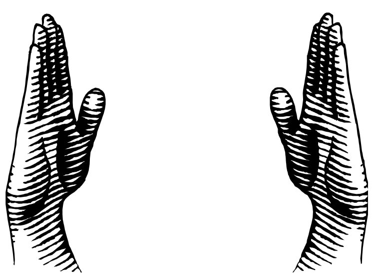
- If ***N*** items can be partitioned into ***k*** cells $A_1, A_2, \ldots, A_k$ with $a_1, a_2, \ldots, a_k$ elements respectively, then the probability distribution of the random variables, $X_1, X_2, \ldots, X_k$ representing the number of elements selected from $A_1, A_2, \ldots, A_k$ in a random sample of size n is:

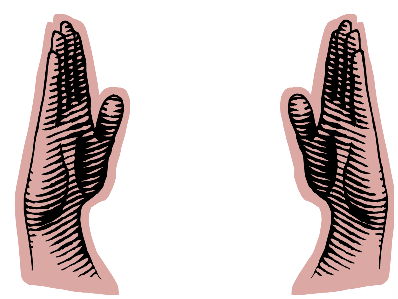
$$f(x_1, x_2 \cdots x_k; a_1, a_2, \cdots a_k, N, n) = \frac{\binom{a_1}{x_1}\binom{a_2}{x_2}\cdots\binom{a_k}{x_k}}{\binom{N}{n}}$$

$$\sum_{i=1}^{k} x_i = n \qquad \sum_{i=1}^{k} a_i = N$$

Example #9

- The probability of a component produced by a certain machine being too small is 10%, too big is 15% and within the accepted dimensions is 75%.
- If seven components are selected randomly from a box that contains 100 components produced by that machine, what is the probability of having four within the accepted dimensions, one too small and two too big?





Example #9 (Sol.)

$$f(x_1,x_2,x_3;a_1,a_2,a_3,N,n)=\frac{\binom{a_1}{x_1}\binom{a_2}{x_2}\binom{a_3}{x_3}}{\binom{N}{n}}$$

N = 100 n = 7

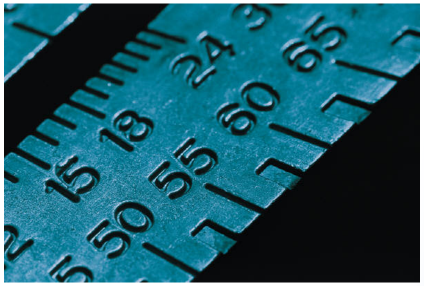
a_1 = 0.75*100 = 75 x_1 = 4

a_2 = 0.10*100 = 10 x_2 = 1

a_3 = 0.15*100 = 15 x_3 = 2

$$f(4,1,2;75,10,15,100,7)=\frac{\binom{75}{4}\binom{10}{1}\binom{15}{2}}{\binom{100}{7}}=\frac{\frac{75!}{4!.71!}*\frac{10!}{1!.9!}*\frac{15!}{2!.13!}}{\frac{100!}{7!.93!}}$$

$$=\frac{\frac{75*74*73*72}{4*3*2*1}*\frac{10}{1}*\frac{15*14}{2*1}}{\frac{100*99*98*97*96*95*94}{7*6*5*4*3*2*1}}=0.0797\cong 0.08$$



Poisson Distribution

- The probability distribution of the *Poisson* random variable X, representing the number of outcomes occurring in a ***given time interval*** or ***specified region*** denoted by ***t***, is:

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

- where λ is the average number of outcomes per unit time, distance, area or volume, and *e* = 2.718281828…..

- Both the mean μ_x and variance σ^2_x of the *Poisson* distribution p(x;λt) are: λ**t**.

Poisson Sums (Table A.2)

- Frequently, we are interested in problems involving the computation of $P(X < r)$ or $P(a \le X \le b)$.

- For this purpose, we have some of the Poisson sums available (Table A.2 in your textbook), where:

$$P(r;\lambda t) = \sum_{x=0}^{r} p\left(x;\lambda t\right)$$

- The sums are available for a range of λt (i.e. μ) =0.10 - 18, and some corresponding values of r.

Example #10

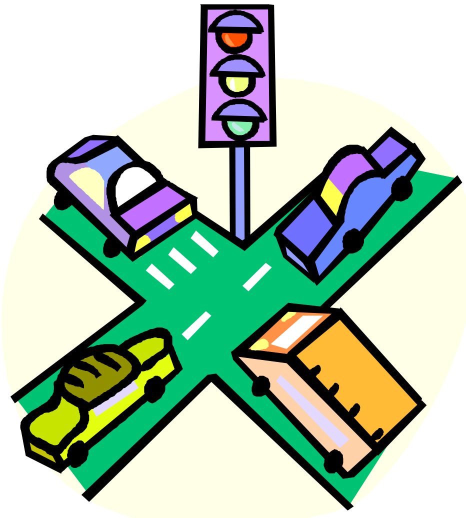
(Ex. 5.58 Textbook):

- On average, a certain intersection results in three traffic accidents per month. What is the probability that for any given month at this intersection:

(a) exactly 5 accidents will occur?

(a) less than 3 accidents will occur?

(b) at least 2 accidents will occur?



Example #10 (Sol.)

$$p(x;\lambda t)=\frac{e^{-\lambda t}(\lambda t)^{x}}{x!}$$

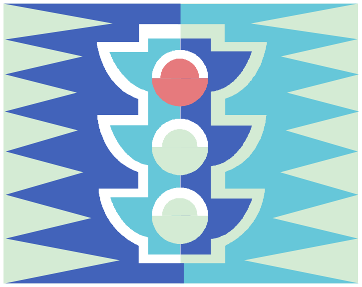
$\lambda t = \mu = 3$

(a) $x = 5$

$$\Rightarrow \quad p(5;3)=\frac{e^{-3}(3)^{5}}{5!}=0.10085 \cong 0.101$$

Example #10 (Sol.)

(a) Or: We can use Table A.2



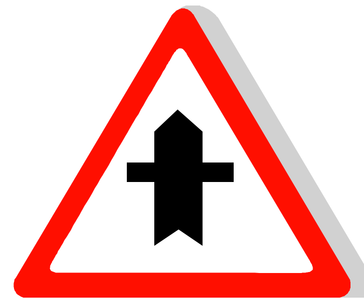
$$P(5;3)=\sum_{x=0}^{5} p(x;\lambda t)-\sum_{x=0}^{4} p(x;\lambda t)$$

r	2.2	2.4	2.6	2.8	3.0	3.5	4.0	4.5	5.0	5.5	6.0
0	0.111	0.091	0.074	0.061	0.050	0.030	0.018	0.011	0.007	0.004	0.002
1	0.355	0.308	0.267	0.231	0.199	0.136	0.092	0.061	0.040	0.027	0.017
2	0.623	0.570	0.518	0.469	0.423	0.321	0.238	0.174	0.125	0.088	0.062
3	0.819	0.779	0.736	0.692	0.647	0.537	0.433	0.342	0.265	0.202	0.151
4	0.928	0.904	0.877	0.848	0.815	0.725	0.629	0.532	0.440	0.358	0.285
5	0.975	0.964	0.951	0.935	0.916	0.858	0.785	0.703	0.616	0.529	0.446
6	0.993	0.988	0.983	0.976	0.966	0.935	0.889	0.831	0.762	0.686	0.606
7	0.998	0.997	0.995	0.992	0.988	0.973	0.949	0.913	0.867	0.809	0.744

$$P(X=5)=0.916-0.815=0.101$$

Example #10 (Sol.)

(b) $P(X < 3) = ?$



$$P(X < 3) = P(X \le 2) = \sum_{x=0}^{2} p(x; \lambda t)$$

r	2.2	2.4	2.6	2.8	3.0	3.5	4.0	4.5	5.0	5.5	6.0
0	0.111	0.091	0.074	0.061	0.050	0.030	0.018	0.011	0.007	0.004	0.002
1	0.355	0.308	0.267	0.231	0.199	0.136	0.092	0.061	0.040	0.027	0.017
2	0.623	0.570	0.518	0.469	0.423	0.321	0.238	0.174	0.125	0.088	0.062
3	0.819	0.779	0.736	0.692	0.647	0.537	0.433	0.342	0.265	0.202	0.151
4	0.928	0.904	0.877	0.848	0.815	0.725	0.629	0.532	0.440	0.358	0.285
5	0.975	0.964	0.951	0.935	0.916	0.858	0.785	0.703	0.616	0.529	0.446
6	0.993	0.988	0.983	0.976	0.966	0.935	0.889	0.831	0.762	0.686	0.606
7	0.998	0.997	0.995	0.992	0.988	0.973	0.949	0.913	0.867	0.809	0.744

$$P(X < 3) = 0.423$$

Example #10 (Sol.)

(c) $P(X \geq 2) = ?$

$$P(X \geq 2) = 1 - P(X \leq 1) = 1 - \sum_{x=0}^{1} p(x; \lambda t)$$

r	2.2	2.4	2.6	2.8	3.0	3.5	4.0	4.5	5.0	5.5	6.0
0	0.111	0.091	0.074	0.061	0.050	0.030	0.018	0.011	0.007	0.004	0.002
1	0.355	0.308	0.267	0.231	0.199	0.136	0.092	0.061	0.040	0.027	0.017
2	0.623	0.570	0.518	0.469	0.423	0.321	0.238	0.174	0.125	0.088	0.062
3	0.819	0.779	0.736	0.692	0.647	0.537	0.433	0.342	0.265	0.202	0.151
4	0.928	0.904	0.877	0.848	0.815	0.725	0.629	0.532	0.440	0.358	0.285
5	0.975	0.964	0.951	0.935	0.916	0.858	0.785	0.703	0.616	0.529	0.446
6	0.993	0.988	0.983	0.976	0.966	0.935	0.889	0.831	0.762	0.686	0.606
7	0.998	0.997	0.995	0.992	0.988	0.973	0.949	0.913	0.867	0.809	0.744

$$\mathrm{P}(\mathrm{X} \geq 2) = 1 - 0.199 = 0.801$$

Textbook Sections

- 5.1
- 5.2
- 5.3
- 5.5
- 5.6