

Example 7.01:

A motor is accepted if it starts at least twice in the first three attempts. On any one attempt, the probability of a success is .8, independently of all other trials. Find the probability that the motor is accepted.

Let $S_n =$ attempt n is successful
 $F_n =$ attempt n is unsuccessful ($= \tilde{S}_n$)
 $E =$ the motor is accepted

then $E = S_1S_2 \cup S_1F_2S_3 \cup F_1S_2S_3$

$$\begin{aligned} \text{so } P[E] &= P[S_1S_2] + P[S_1F_2S_3] + P[F_1S_2S_3] && \text{[mutually exclusive events]} \\ &= (.8)^2 + (.2)(.8)^2 + (.2)(.8)^2 && \text{[independent events]} \\ &= .64 \times (1 + .2 + .2) = \underline{\underline{.896}} \end{aligned}$$

[Note: if $X =$ (number of attempts needed to obtain the second success), then X follows a negative binomial distribution $nb(x; 2, .8)$ (Devore, section 3.5, pages 118-120).]

Example 7.02:

A Poisson process [textbook, section 3.6, pages 121-124] has a probability mass function defined by

$$f(x) = P[X = x] = \frac{e^{-\mu} \mu^x}{x!} \quad (x = 0, 1, 2, \dots)$$

where $\mu = E[X]$.

(a) Show that $f(x)$ is a valid probability mass function.

- ▶ $\frac{e^{-\mu} \mu^x}{x!} > 0$ for all $x \in \mathbb{N}$ and $x = 0$. Therefore $f(x) \geq 0$.
- ▶ $\sum_{\forall x} f(x) = \sum_{x=0}^{\infty} \frac{e^{-\mu} \mu^x}{x!} = e^{-\mu} \sum_{x=0}^{\infty} \frac{\mu^x}{x!}$

Example 7.02(a) (continued)

$$\text{But } \sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x$$

$$\text{So } \sum_{\forall x} f(x) = e^{-\mu} \times e^{\mu} = 1$$

Therefore $f(x)$ is a valid probability mass function.

In a factory producing insulated cables, the number of cracks in the cables may be modelled to a good approximation as a Poisson process with a mean rate of occurrence of 3.4 cracks per metre. Find the probability that

- (b) there are less than 20 cracks in a five metre long cable.
- (c) there are exactly 20 cracks in a five metre long cable.

$$(b) \quad \mu = 3.4 \text{ (cracks per metre)} \times 5 \text{ (metres)} = 17$$

$$P[X < 20] = \sum_{x=0}^{19} e^{-17} \frac{17^x}{x!} = e^{-17} \left(1 + \frac{17}{1} + \frac{17 \cdot 17}{1 \cdot 2} + \frac{17 \cdot 17 \cdot 17}{1 \cdot 2 \cdot 3} + \dots + \frac{17^{19}}{19!} \right)$$

By using a computer package (or by direct calculation), $P[X < 20] = .736\ 321\ 7\dots$

Therefore $P[\text{less than 20 cracks in a five metre long cable}] \approx \underline{\underline{.7363}}$

$$(c) \quad P[X = 20] = \frac{e^{-17} 17^{20}}{20!} \approx \underline{\underline{.0692}}$$

Every twentieth crack, on average, is serious enough to cause failure of the system to which it is connected. Find the probabilities of

- (d) no serious cracks in a five metre long cable.
- (e) less than three serious cracks in a five metre long cable.

To find the probability values in example 7.02, the following extracts from the Excel spreadsheet file "www.engr.mun.ca/~ggeorge/3423/demos/Poisson.xls" can be used:

The value of the parameter μ of the Poisson distribution is $\mu = 17$ in parts (b) and (c).

In part (b), $P[X < 20] = P[X \leq 19]$, so find the value of the Poisson c.d.f. $F(19; 17)$.

In part (c), find $P[X = 20]$, so find the value of the Poisson p.m.f. $f(20; 17)$.

$$\mu = 17$$

p.m.f.		c.d.f.	
x	$P[X = x]$	x	$P[X \leq x]$
10	0.02300	10	0.04912
11	0.03554	11	0.08467
12	0.05036	12	0.13502
13	0.06585	13	0.20087
14	0.07996	14	0.28083
15	0.09062	15	0.37145
16	0.09628	16	0.46774
17	0.09628	17	0.56402
18	0.09094	18	0.65496
19	0.08136	19	0.73632
20	0.06916	20	0.80548
21	0.05599	21	0.86147
22	0.04326	22	0.90473
23	0.03198	23	0.93670
24	0.02265	24	0.95935
25	0.01540	25	0.97476

(d) Every twentieth defective crack is serious $\Rightarrow \mu \rightarrow 17/20 = 0.85$

$$P[X=0] = \frac{e^{-0.85}(0.85)^0}{0!} \approx \underline{\underline{.4274}}$$

$$(e) \quad P[X < 3] = \sum_{x=0}^2 \frac{e^{-0.85}(0.85)^x}{x!} = e^{-0.85} \left(1 + \frac{0.85}{1} + \frac{(0.85)^2}{2} \right) \approx \underline{\underline{.9451}}$$

[Note that the Poisson c.d.f. is tabulated for only a few selected values of μ and x in Table A.2 of Devore. The table below is based upon the Excel spreadsheet at www.engr.mun.ca/~ggeorge/3423/demos/Poisson.xls.]

The value of the parameter μ of the Poisson distribution is $\mu = 0.85$ in parts (d) and (e). In part (d), find $P[X=0]$, so find the value of the Poisson p.m.f. $f(0; 0.85)$. In part (e), $P[X < 3] = P[X \leq 2]$, so find the value of the Poisson c.d.f. $F(2; 0.85)$.

$\mu = 0.85$

p.m.f.		c.d.f.	
x	$P[X=x]$	x	$P[X \leq x]$
-2	0.00000	-2	0.00000
-1	0.00000	-1	0.00000
0	0.42741	0	0.42741
1	0.36330	1	0.79072
2	0.15440	2	0.94512
3	0.04375	3	0.98887
4	0.00930	4	0.99817
5	0.00158	5	0.99975
6	0.00022	6	0.99997
7	0.00003	7	1.00000
8	0.00000	8	1.00000
9	0.00000	9	1.00000
10	0.00000	10	1.00000

We can also verify that, for *any* Poisson process,

$$\boxed{E[X] = V[X] = \mu}$$

Proof (non-examinable):

$$E[X] = \sum_{x=0}^{\infty} x \left(\frac{e^{-\mu} \mu^x}{x!} \right) = 0 + \sum_{x=1}^{\infty} \left(\frac{e^{-\mu} \mu^x}{(x-1)!} \right) \quad \left(\because \frac{x!}{x} = (x-1)! \right)$$

Let $y = x - 1$, then

$$E[X] = e^{-\mu} \sum_{y=0}^{\infty} \frac{\mu^{y+1}}{y!} = \mu e^{-\mu} \sum_{y=0}^{\infty} \frac{\mu^y}{y!} = \mu e^{-\mu} e^{\mu}$$

Therefore $E[X] = \mu$.

$$\begin{aligned} E[X^2] &= \sum_{x=0}^{\infty} x^2 \left(\frac{e^{-\mu} \mu^x}{x!} \right) = 0 + \sum_{x=1}^{\infty} \left(\frac{e^{-\mu} \mu^x x}{(x-1)!} \right) \\ &= \sum_{y=0}^{\infty} \frac{e^{-\mu} \mu^{y+1} (y+1)}{y!} = \mu \sum_{y=0}^{\infty} \frac{y e^{-\mu} \mu^y}{y!} + \mu e^{-\mu} \sum_{y=0}^{\infty} \frac{\mu^y}{y!} \\ &= \mu E[X] + \mu e^{-\mu} e^{\mu} = \mu^2 + \mu \end{aligned}$$

But $V[X] = E[X^2] - (E[X])^2$.

Therefore $V[X] = \mu^2 + \mu - \mu^2 = \mu$ (and $\sigma = \sqrt{\mu}$).

The Binomial Distribution

Example 7.03

It is known that 60% of all components emerging from an initial production process are good and 40% are defective. A random sample of ten components is drawn.

- (a) Find the probability that exactly six components in the random sample are good.

In a single trial (a single component in the random sample) let

S = success (the component is good) and

F = failure (the component is defective) ($= \tilde{S}$)

Also let

E = the desired event (six good and four defective components in the random sample)

Then one way in which the desired event can occur is if the first six components tested are all good and the remaining four are all defective. The probability of this outcome (SSSSSSFFFF in that order) is

$$P[S^6F^4] = (.60)^6(.40)^4 = .00119\dots$$

The number of distinct rearrangements of six successes in ten trials is

$${}^{10}C_6 = \binom{10}{6} = \frac{10 \times 9 \times 8 \times 7}{4 \times 3 \times 2 \times 1} = 210$$

Therefore $P[E] = 210 \times .00119\dots \approx \underline{\underline{.251}}$

- (b) Find the probability mass function for X = (the number of good components in the random sample).

$$P[S^x F^{10-x}] = (.60)^x (.40)^{10-x}$$

But these x successes can be rearranged among the 10 trials in ${}^{10}C_x$ distinct ways.

Therefore the probability mass function of X is

$$P[\text{exactly } x \text{ successes}] = P[X=x] = {}^{10}C_x (.60)^x (.40)^{10-x} \quad (x = 0, 1, 2, \dots, 10)$$

In general, if the random quantity X represents the number of successes in n trials then X will have a binomial probability distribution if and only if the following conditions hold:

- (1) Each trial has exactly two complementary outcomes (“success” and “failure”);
- (2) The probability of success is constant across all trials;
- (3) The outcome of each trial is independent of all other trials;
- (4) The sample size n is fixed.

The binomial probability mass function is

$$P[X = x] = b(x; n, p) = \binom{n}{x} p^x (1-p)^{n-x}$$

Condition (3) holds only if the sampling is either with replacement or from an infinite population (or both). If the sampling is without replacement from a finite but large population, then condition (3) may be approximately true.

Example 7.04

Show that condition (3) is not satisfied if a random sample of size 2 is taken from a population of size 5 with two successes. Show that condition (3) is almost satisfied if a random sample of size 2 is taken from a population of size 5000 with 2000 successes.

S S F F F

Let S_1 = success on trial 1
and S_2 = success on trial 2

then $P[S_2] = 2/5 = .400\ 00$
but $P[S_2 | S_1] = 1/4 = .250\ 00$

Therefore the outcomes of the two trials are not independent.
[The exact probability distribution is hypergeometric.]

2,000 S's 3,000 F's

Let S_1 = success on trial 1
and S_2 = success on trial 2

then $P[S_2] = 2,000 / 5,000 = .400\ 00$

but $P[S_2 | S_1] = \frac{1,999}{4,999} \approx .399\ 88$

Therefore the outcomes of the two trials are nearly independent.

The binomial cumulative distribution function is

$$\begin{aligned}
 B(x; n, p) &= \sum_{y=0}^x b(y; n, p) \\
 &= b(0; n, p) + b(1; n, p) + b(2; n, p) + \dots + b(x; n, p)
 \end{aligned}$$

The binomial *cdf* $B(x; n, p)$ is tabulated in Devore, table A.1, for $n = 5, 10, 15, 20, 25$ and a few values of p . It can be evaluated for any valid choice of (n, p) using the file at "www.engr.mun.ca/~ggeorge/3423/demos/Binomial.xls".

Example 7.05

The probability mass function of the random quantity X is known to be binomial with parameters $n = 10$ and $p = .60$. Find $P[4 < X < 8]$.

$$\begin{aligned}
 P[4 < X < 8] &= P[4 < X \leq 7] \\
 &= B(7; 10, .60) - B(4; 10, .60) \\
 &= .833 - .166 \quad \text{(using Table A.1)} \\
 &= .667
 \end{aligned}$$

or, using the Web file "Binomial.xls",

$$\begin{aligned}
 B(7; 10, .60) - B(4; 10, .60) &= .83271 - .16624 \\
 &\approx \underline{\underline{.6665}}
 \end{aligned}$$

OR

$$\begin{aligned}
 P[4 < X < 8] &= \\
 &P[X = 5] + P[X = 6] + P[X = 7] \\
 &= {}^{10}C_5 (.60)^5 (.40)^5 \\
 &+ {}^{10}C_6 (.60)^6 (.40)^4 \\
 &+ {}^{10}C_7 (.60)^7 (.40)^3 \\
 &\approx \underline{\underline{.666472}}
 \end{aligned}$$

[A spreadsheet can be used to carry out these calculations. MINITAB and Excel both contain the binomial *pmf* and *cdf*, as does the Excel file at

"www.engr.mun.ca/~ggeorge/3423/demos/Binomial.xls".]

To find the values of $B(7; 10, .60)$ and $B(4; 10, .60)$, the following partial table of values for $b(x; 10, .60)$ and $B(x; 10, .60)$, (drawn from the Excel file), may be used:

$n = 10$		$p = 0.6$	
p.m.f.		c.d.f.	
x	$P[X = x]$	x	$P[X \leq x]$
0	0.00010	0	0.00010
1	0.00157	1	0.00168
2	0.01062	2	0.01229
3	0.04247	3	0.05476
4	0.11148	4	0.16624
5	0.20066	5	0.36690
6	0.25082	6	0.61772
7	0.21499	7	0.83271
8	0.12093	8	0.95364
9	0.04031	9	0.99395
10	0.00605	10	1.00000

Example 7.06

Ten per cent of all items in a large production run are known to be defective. A random sample of 20 items is drawn.

- (a) Prove that the random quantity X = (the number of defective items in the random sample) has a binomial probability mass function.
- (b) Find the probability that more than two items in the random sample are defective.
- (c) How many defective items does one expect to find in the random sample?

- (a) Let X = the number of defective items in the random sample.
- ▶ each trial (item) has a complementary pair of outcomes (defective = “success”, good = “failure”)
 - ▶ $P[\text{success}] = \text{constant} = 10\%$
 - ▶ Trials are independent to a good approximation (because the random sample is drawn from a *large* population)
 - ▶ $n = 20$ is fixed.

All four conditions are satisfied. Therefore X does follow a binomial probability distribution. $P[X = x] = b(x; 20, .10)$.

(b) $P[X > 2] = 1 - P[X \leq 2]$

$$= 1 - B(2; 20, .10)$$

$$= 1 - .67693 \approx \underline{\underline{.3231}}$$

OR (by direct calculation of the p.m.f. values),

$$1 - P[X \leq 2] = 1 - (P[X = 0] + P[X = 1] + P[X = 2])$$

$$= 1 - \{ (.9)^{20} + 20(.9)^{19}(.1)^1 + 190(.9)^{18}(.1)^2 \}$$

$$= 1 - (.121... + .270... + .285...)$$

$$\approx \underline{\underline{.323073}}$$

Side note:

$$P[X > 2] = P[2 < X \leq 20] \\ = B(20; 20, .1) - B(2; 20, .1)$$

But $B(n; n, p) = 1$ for any n : it absolutely certain to get at most n successes in n attempts. Therefore

$$P[X > 2] = 1 - B(2; 20, .1)$$

(as in the main solution)

To find the value of $B(2; 20, .10)$, the following extracts from the Excel spreadsheet file "www.engr.mun.ca/~ggeorge/3423/demos/Binomial.xls" can be used:

Enter values in the three highlighted boxes:

Number of trials

$$n = \boxed{20} \leftarrow \text{must be } > 0$$

Prob. success in a trial

$$p = \boxed{0.1} \leftarrow \text{must be in } [0, 1]$$

Prob. failure in a trial

$$q = 0.9$$

Mean = expected value

$$E[X] = \mu = 2$$

Variance

$$V[X] = \sigma^2 = 1.8$$

Standard deviation

$$\sigma = 1.34164$$

Enter a value for X :

$$x = \boxed{2} \leftarrow \text{must be in } [0, n]$$

Binomial p.m.f.:

$$b(x; n, p) = \boxed{0.28518}$$

Binomial c.d.f.:

$$B(x; n, p) = \boxed{0.67693}$$

OR

the partial table of values for $b(x; 20, .10)$ and $B(x; 20, .10)$ may be used instead:

$n = 20$		$p = 0.1$	
	p.m.f.		c.d.f.
x	$P[X = x]$	x	$P[X \leq x]$
0	0.12158	0	0.12158
1	0.27017	1	0.39175
2	0.28518	2	0.67693
3	0.19012	3	0.86705
4	0.08978	4	0.95683
5	0.03192	5	0.98875
6	0.00887	6	0.99761
7	0.00197	7	0.99958
8	0.00036	8	0.99994
9	0.00005	9	0.99999
10	0.00001	10	1.00000

(c) For any binomial random quantity X ,

$$E[X] = np \quad V[X] = np(1-p)$$

Thus $\mu = E[X] = 20 \times .10 = \underline{2}$

The proof of the formulae for the mean and variance of the binomial random quantity is *not* examinable, but is presented here:

$$\begin{aligned} \mu &= E[X] = \sum_{x=0}^n x \cdot b(x; n, p) = \sum_{x=0}^n x \cdot \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\ &= 0 + \sum_{x=1}^n \frac{n(n-1)!}{(x-1)!((n-1)-(x-1))!} p p^{x-1} (1-p)^{((n-1)-(x-1))} \end{aligned}$$

Let $y = x - 1$ and $m = n - 1$ then

$$E[X] = np \sum_{y=0}^m \frac{m!}{y!(m-y)!} p^y (1-p)^{m-y} = np \sum_{y=0}^m b(y; m, p) = np \times 1 = \underline{np}$$

$$\begin{aligned} E[X^2] &= \sum_{x=0}^n x^2 \cdot b(x; n, p) = \sum_{x=0}^n x^2 \cdot \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\ &= 0 + \sum_{x=1}^n x \cdot \frac{n!}{(x-1)!(n-x)!} p^x (1-p)^{n-x} = \sum_{x=1}^n (x-1+1) \cdot \frac{n!}{(x-1)!(n-x)!} p^x (1-p)^{n-x} \\ &= 0 + \sum_{x=2}^n \frac{n!}{(x-2)!(n-x)!} p^x (1-p)^{n-x} + \sum_{x=1}^n \frac{n!}{(x-1)!(n-x)!} p^x (1-p)^{n-x} \end{aligned}$$

But the second of these summations is, from the derivation of $E[X]$ above, equal to $E[X]$.

Therefore

$$E[X^2] = n(n-1) p^2 \sum_{x=2}^n \frac{(n-2)!}{(x-2)!((n-2)-(x-2))!} p^{x-2} (1-p)^{((n-2)-(x-2))} + \mu$$

Let $y = x - 2$ and $m = n - 2$ then

$$\begin{aligned} E[X^2] &= n(n-1) p^2 \sum_{y=0}^m \frac{m!}{y!(m-y)!} p^y (1-p)^{m-y} + \mu \\ &= n(n-1) p^2 \sum_{y=0}^m b(y; m, p) + \mu = n(n-1) p^2 + \mu \end{aligned}$$

$$\begin{aligned} \text{Therefore } V[X] &= E[X^2] - (E[X])^2 \\ &= n(n-1)p^2 + np - n^2p^2 = n^2p^2 - np^2 + np - n^2p^2 = \underline{np(1-p)} \end{aligned}$$

Example 7.07

Find the probability of (a) exactly one '5' or '6' (b) at least one '5' or '6' when four dice are thrown.

(c) How many times, on average, do you expect a '5' or '6' to occur when four dice are thrown?

Let $X =$ (number of times a '5' or '6' occurs), then $P[X = x] = b(x; 4, 1/3)$.

[One can check that all four conditions for a p.m.f. to be binomial are valid.]

(a) $P[X = 1] = b(1; 4, 1/3) =$

$${}^4C_1 \left(\frac{1}{3}\right)^1 \left(\frac{2}{3}\right)^3 = \frac{32}{81} = .395061\dots \approx \underline{\underline{.395}}$$

(b) $P[X \geq 1] = 1 - P[X < 1] = 1 - P[X = 0] =$

$$1 - \left(\frac{2}{3}\right)^4 = \frac{65}{81} = .802469\dots \approx \underline{\underline{.802}}$$

(c) $E[X] = np = 4/3 \approx \underline{\underline{1.333}}$

Note that the limit of the binomial probability mass function as $n \rightarrow \infty$ and $p \rightarrow 0$ in such a way that μ is constant is the Poisson pmf:

$b(x; n, p) \rightarrow \text{Poisson}(x; \mu) \text{ as } n \rightarrow \infty \text{ with } \mu = \text{constant}$
--

Proof (*non-examinable*):

$$b(x; n, p) = \frac{n!}{x! (n-x)!} p^x (1-p)^{n-x}. \text{ But } \mu = np \Rightarrow p = \mu/n, \text{ so}$$

$$b(x; n, p) = \frac{n(n-1)(n-2)\cdots(n-x+1)(n-x)!}{x!(n-x)!} \frac{\mu^x}{n^x} \left(1 - \frac{\mu}{n}\right)^{n-x}$$

$$= \frac{\mu^x}{x!} \left(1 - \frac{\mu}{n}\right)^{-x} \left(1 - \frac{\mu}{n}\right)^n \frac{n}{n} \cdot \frac{n-1}{n} \cdot \frac{n-2}{n} \cdots \frac{n-x+1}{n}$$

$$= \frac{\mu^x}{x!} \left(1 - \frac{\mu}{n}\right)^{-x} \left(1 - \frac{\mu}{n}\right)^n 1 \cdot \left(1 - \frac{1}{n}\right) \cdot \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{x-1}{n}\right)$$

$$\Rightarrow \lim_{\substack{p \rightarrow 0 \\ np = \mu}} b(x; n, p) = \frac{\mu^x}{x!} (1-0)^{-x} (1-0)^x \lim_{n \rightarrow \infty} \left(1 - \frac{\mu}{n}\right)^n = \frac{\mu^x e^{-\mu}}{x!} = \text{Poisson}(x; \mu)$$

Note also that $V[X] = np(1-p) \rightarrow \mu(1-0) = \mu$.

Example 7.08

A production process in a factory has a defect rate of 2%. What is the smallest sample size for which the probability of encountering at least one defective item exceeds 95%?

In other words, find the least n such that $P[X > 0] > 0.95$, when $P[X = x] = b(x; n, .02)$.

$$\begin{aligned} P[X > 0] &= 1 - P[X \leq 0] \\ &= 1 - P[X = 0] = \\ &= 1 - (.98)^n \end{aligned}$$

But we require $P[X > 0] > .95$

$$\therefore 1 - (.98)^n > .95$$

$$\Rightarrow 1 - .95 > (.98)^n$$

$$\Rightarrow \ln(1 - .95) > \ln((.98)^n) \quad \text{['ln' is a monotonic function]}$$

$$\Rightarrow \ln(.05) > n \ln(.98)$$

$$\Rightarrow n > \frac{\ln(.05)}{\ln(.98)} = \frac{-2.9957\dots}{-0.0202\dots} \quad \text{[sign reverses } \therefore \text{ negative divisor]}$$

$$\Rightarrow n > 148.28\dots$$

$$\therefore n_{\min} = \underline{149} \quad \text{[not 148, as can be seen below.]}$$

As a check, $n = 148 \Rightarrow P[X > 0] = 1 - (.98)^{148} \approx 94.97\%$
 $n = 149 \Rightarrow P[X > 0] = 1 - (.98)^{149} \approx 95.07\%$.
