Binomial Distribution (7 pages; 16/2/17)

(1) $X \sim B(n, p)$ (discrete random variable)

$\Rightarrow P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \quad x = 0,1,2, \ \ldots, n$

$= 0$ otherwise

Notes

(i) The 'Binomial coefficient', $\binom{n}{x} = \frac{n!}{x!(n-x)!}$ can also be written as ${}^nC_r$

(ii) $1 - p$ is often written as q

(iii) See Appendix B for a demonstration that $\sum P(X = x) = 1$ (as is necessary for a probability distribution).

Example: A factory produces computer laptops. The probability of a laptop working properly is $p = 0.6$ ("probability of success").

If there are $n = 5$ laptops coming off the production line, and X is the number of working laptops, then $X \sim B(5, \ 0.6)$ and the probability of at least one laptop working properly is

$1 - P(X = 0) = 1 - \binom{5}{0}(0.6)^0(1 - 0.6)^{5-0}$

$= 1 - (0.4)^5 = 0.98976 = 0.990$ (3sf)

(2) Derivation of the Binomial probability

Consider $P(X = 3)$ in the above example.

For one particular ordering of the successes and failures; say $SSFSF$,

$P(SSFSF) = (0.6)(0.6)(0.4)(0.6)(0.4) = (0.6)^3(0.4)^2$

The possible orderings are:

SSSFF, SSFSF, SSFFS, SFSSF, SFSFS,

SFFSS, FSSSF, FSSFS, FSFSS, FFSSS

This is the number of ways of choosing 3 positions for S, out of the total of 5; ie $\binom{5}{3} = \frac{5!}{3!2!}$ (see Appendix A, for the derivation of this).

Each ordering is equally likely, so that

$P(X = 3) = \binom{5}{3}(0.6)^3(0.4)^2$

(3) Conditions that need to apply in order for the Binomial model to be valid

(i) The outcomes of the n trials must be random and independent of each other.

(ii) The probability of success must be constant over the n trials.

(4) Cumulative tables

See Appendix C.

To avoid manual calculations, note that

$P(X = 3) = P(X \leq 3) - P(X \leq 2)$

(5) Mean of a Binomial Variable

If $X \sim B(n, p)$, $E(X) = \sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x} x$

$= \sum_{x=1}^{n} \binom{n}{x} p^x (1-p)^{n-x} x$

$= \sum_{x=1}^{n} \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} x$

$= np \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!(n-x)!} p^{(x-1)} (1-p)^{n-x}$

$= np \sum_{x-1=0}^{n-1} \frac{(n-1)!}{(x-1)!(n-x)!} p^{(x-1)} (1-p)^{n-x}$

Let $u = x - 1$ and $N = n - 1$

Then $E(X) = np \sum_{u=0}^{N} \frac{N!}{u!(N-u)!} p^u (1-p)^{N-u}$

$= np \sum_u P(X = u) = np$

(6) Variance of a Binomial Variable

$Var(X) = E(X^2) - \mu^2 = \mathrm{E}[\mathrm{X}(\mathrm{X}-1)+\mathrm{X}] - \mu^2 = \mathrm{E}[\mathrm{X}(\mathrm{X}-1)] + \mu - \mu^2$

$= [\sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x} x(x-1)] + np - (np)^2$

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

$= n(n-1)p^2 \sum_{x=2}^{n} \frac{(n-2)!}{(x-2)!(n-x)!} p^{(x-2)} (1-p)^{n-x}] + np - (np)^2$

Let $u = x - 2$ and $N = n - 2$

Then Var(X)

$= n(n-1)p^2 \sum_{u=0}^{N} \frac{N!}{u!(N-u)!} p^u (1-p)^{N-u}] + np - (np)^2$

$= n(n-1)p^2 \sum_u P(X = u) + np - (np)^2 \quad = n(n-1)p^2 + np - (np)^2$

$= np\{(n-1)p + 1 - np\} = np(1-p)$

(7) Approximations to the Binomial distribution

For large n and small p, the Binomial distribution can be approximated by the Poisson distribution. For large n and moderate p, the Binomial distribution can be approximated by the Normal distribution (though a smaller value of p can be tolerated if n is large enough).

See "Approximations to the Binomial and Poisson Distributions".

(8) Miscellaneous

(i) n is occasionally referred to as the index, and p as the parameter.

Appendix A: Derivation of $\binom{n}{r}$: the number of ways of choosing r items from n

Example of $\binom{5}{3}$

There are $5 \times 4 \times 3 \times 2 \times 1 = 5!$ different orderings of ABCDE.

[There are 5 choices for the 1st position; then for each of these there are 4 choices for the 2nd position etc.]

Now consider the number of different orderings of ABCCC.

The following all count as the same:

$BC_1C_2AC_3, BC_1C_3AC_2, BC_2C_1AC_3, BC_2C_3AC_1, BC_3C_1AC_2$, $\mathrm{B}C_3C_2AC_1$

and similarly for $CACBC\ etc$

Thus there is a 3! duplication of the Cs.

So there are $\frac{5!}{3!}$ different orderings of ABCCC

Now consider the number of different orderings of BBCCC.

In the same way as above, there is a 2! duplication of the Bs.

Thus the number of ways of arranging BBCCC is $\frac{5!}{3!2!}$

Notes

(i) It can be shown that $\binom{n}{r}$ is the rth value in the nth row of Pascal's triangle (where r starts at 0, and the nth row starts $1, n, \ldots$)

(ii) $\binom{5}{3} = \frac{5!}{3!2!} = \binom{5}{2}$, as the number of ways of choosing 3 positions for S (in the example used above) is the same as the number of ways of choosing 2 positions for F.

Consider also the symmetry of Pascal's triangle.

(iii) $\binom{n}{r} = \frac{n!}{r!(n-r)!} = \frac{n(n-1)\ldots[r\ items]}{r!}$

Thus, $\binom{20}{17} = \binom{20}{3} = \frac{20(19)(18)}{3!}$

Appendix B: Use of $\binom{n}{r}$ in the Binomial expansion

To show that $(a+b)^n = \sum_{r=0}^{n} \binom{n}{r} a^r b^{n-r}$:

eg $(a+b)^5 = (a+b)(a+b)(a+b)(a+b)(a+b)$

The number of times that a^3b^2 appears in the expansion of this expression is the number of ways in which we can choose 3 out of the 5 brackets for the $a's$ (with remaining 2 brackets giving the $b's$); ie $\binom{5}{3}$

Note that when $q = 1-p$, $(p+q)^n = \sum_{r=0}^{n} \binom{n}{r} p^r q^{n-r}$, but $(p+q)^n = 1^n = 1$

Thus, the Binomial probabilities add up to 1, as expected.

Appendix C: Cumulative tables

BINOMIAL CUMULATIVE DISTRIBUTION FUNCTION

The tabulated value is $P(X \le x)$, where X has a binomial distribution with index n and parameter p.

$p =$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
$n = 5, x = 0$	0.7738	0.5905	0.4437	0.3277	0.2373	0.1681	0.1160	0.0778	0.0503	0.0312
1	0.9774	0.9185	0.8352	0.7373	0.6328	0.5282	0.4284	0.3370	0.2562	0.1875
2	0.9988	0.9914	0.9734	0.9421	0.8965	0.8369	0.7648	0.6826	0.5931	0.5000
3	1.0000	0.9995	0.9978	0.9933	0.9844	0.9692	0.9460	0.9130	0.8688	0.8125
4	1.0000	1.0000	0.9999	0.9997	0.9990	0.9976	0.9947	0.9898	0.9815	0.9688
$n = 6, x = 0$	0.7351	0.5314	0.3771	0.2621	0.1780	0.1176	0.0754	0.0467	0.0277	0.0156
1	0.9672	0.8857	0.7765	0.6554	0.5339	0.4202	0.3191	0.2333	0.1636	0.1094
2	0.9978	0.9842	0.9527	0.9011	0.8306	0.7443	0.6471	0.5443	0.4415	0.3438
3	0.9999	0.9987	0.9941	0.9830	0.9624	0.9295	0.8826	0.8208	0.7447	0.6563
4	1.0000	0.9999	0.9996	0.9984	0.9954	0.9891	0.9777	0.9590	0.9308	0.8906
5	1.0000	1.0000	1.0000	0.9999	0.9998	0.9993	0.9982	0.9959	0.9917	0.9844
$n = 7, x = 0$	0.6983	0.4783	0.3206	0.2097	0.1335	0.0824	0.0490	0.0280	0.0152	0.0078
1	0.9556	0.8503	0.7166	0.5767	0.4449	0.3294	0.2338	0.1586	0.1024	0.0625
2	0.9962	0.9743	0.9262	0.8520	0.7564	0.6471	0.5323	0.4199	0.3164	0.2266
3	0.9998	0.9973	0.9879	0.9667	0.9294	0.8740	0.8002	0.7102	0.6083	0.5000
4	1.0000	0.9998	0.9988	0.9953	0.9871	0.9712	0.9444	0.9037	0.8471	0.7734
5	1.0000	1.0000	0.9999	0.9996	0.9987	0.9962	0.9910	0.9812	0.9643	0.9375
6	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998	0.9994	0.9984	0.9963	0.9922
$n = 8, x = 0$	0.6634	0.4305	0.2725	0.1678	0.1001	0.0576	0.0319	0.0168	0.0084	0.0039
1	0.9428	0.8131	0.6572	0.5033	0.3671	0.2553	0.1691	0.1064	0.0632	0.0352
2	0.9942	0.9619	0.8948	0.7969	0.6785	0.5518	0.4278	0.3154	0.2201	0.1445
3	0.9996	0.9950	0.9786	0.9437	0.8862	0.8059	0.7064	0.5941	0.4770	0.3633
4	1.0000	0.9996	0.9971	0.9896	0.9727	0.9420	0.8939	0.8263	0.7396	0.6367
5	1.0000	1.0000	0.9998	0.9988	0.9958	0.9887	0.9747	0.9502	0.9115	0.8555
6	1.0000	1.0000	1.0000	0.9999	0.9996	0.9987	0.9964	0.9915	0.9819	0.9648
7	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998	0.9993	0.9983	0.9961
$n = 9, x = 0$	0.6302	0.3874	0.2316	0.1342	0.0751	0.0404	0.0207	0.0101	0.0046	0.0020
1	0.9288	0.7748	0.5995	0.4362	0.3003	0.1960	0.1211	0.0705	0.0385	0.0195
2	0.9916	0.9470	0.8591	0.7382	0.6007	0.4628	0.3373	0.2318	0.1495	0.0898
3	0.9994	0.9917	0.9661	0.9144	0.8343	0.7297	0.6089	0.4826	0.3614	0.2539
4	1.0000	0.9991	0.9944	0.9804	0.9511	0.9012	0.8283	0.7334	0.6214	0.5000
5	1.0000	0.9999	0.9994	0.9969	0.9900	0.9747	0.9464	0.9006	0.8342	0.7461
6	1.0000	1.0000	1.0000	0.9997	0.9987	0.9957	0.9888	0.9750	0.9502	0.9102
7	1.0000	1.0000	1.0000	1.0000	0.9999	0.9996	0.9986	0.9962	0.9909	0.9805
8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9992	0.9980

$n = 10, x = 0$	0.5987	0.3487	0.1969	0.1074	0.0563	0.0282	0.0135	0.0060	0.0025	0.0010
1	0.9139	0.7361	0.5443	0.3758	0.2440	0.1493	0.0860	0.0464	0.0233	0.0107
2	0.9885	0.9298	0.8202	0.6778	0.5256	0.3828	0.2616	0.1673	0.0996	0.0547
3	0.9990	0.9872	0.9500	0.8791	0.7759	0.6496	0.5138	0.3823	0.2660	0.1719
4	0.9999	0.9984	0.9901	0.9672	0.9219	0.8497	0.7515	0.6331	0.5044	0.3770
5	1.0000	0.9999	0.9986	0.9936	0.9803	0.9527	0.9051	0.8338	0.7384	0.6230
6	1.0000	1.0000	0.9999	0.9991	0.9965	0.9894	0.9740	0.9452	0.8980	0.8281
7	1.0000	1.0000	1.0000	0.9999	0.9996	0.9984	0.9952	0.9877	0.9726	0.9453
8	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9995	0.9983	0.9955	0.9893
9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9990