Binomial Distributions

Conditions for a **Binomial Distribution**: **[Memorize this!]**

- There are *n* trials or repetitions
- There are 2 outcomes for each trial, S or F
- The P(Success), P, for each trial is constant. **Note:** Often, the probability is expressed as a "Proportion" e.g. the true proportion of defectives in a batch of TV–sets in 8% => P = 8% = 0.08 [Here, somewhat amusingly, Success ~ Defective TV–set...but that's OK!]
- The trials are identical and independent.
- X is a r.v. denoting *number* of successes, *x*, in *n* trials [Note: there are n – *x* failures in *n* trials]
- Notation: X ~ B(n, P) Also, Q = P(Failure) = 1 – P

Formulas:

- P(X **=** *x* successes) = Binom**P**df (n, P, *x)*
  $= {}^{n}C_{x}\, S^{x}\, F^{n-x} \sim {}^{n}C_{x}\, P^{x}\, Q^{n-x}$ where Q = P(Failure) = 1 – P
- P(X **≤** *k* successes) = Binom**C**df (n, P, *k*)
  = P(X = 0) + P(X = 1) + P(X = 2) + ... P(X = *k*)
  = Σ [*from x* = 0 *to x* = *k*] $\;{}^{n}C_{x}\, S^{x}\, F^{n-x}$
  = Σ [*from x* = 0 *to x* = *k*] $\;{}^{n}C_{x}\, P^{x}\, (1-P)^{n-x}$
- **Caution:** For P(X **<** *x*), P(X **≥** *x*) and P(X **>** *x*) Qs, since the TI 83s and 84s can perform *only* ≤ operations for Binom**C**df...you must "transform" the Q into something *usable*: if uncertain, draw a number line and use common–sense / logic!

Examples to help you with interpretations:

- P(5 successes) = P(X = 5) = Binom**P**df(n, P, 5)
- P(at most 10 successes) = P(X ≤ 10) = P(X = 0) + P(X = 1) + ...P(X = 10)
  = Binom**C**df(n, P, 10)
- P(fewer than 7 successes) = **P(X < 7)** = P(X = 0) + ...P(X = 6)
  = **P(X ≤ 6)** = Binom**C**df(n, P, 6)
- P(more than 11 successes) = **P(X > 11)**
  = 1 – P(X ≤ 11)
  = 1 – Binom**C**df(n, P, 11)
- P(at least 8 successes) = P(X ≥ 8)
  = 1 – P(X ≤ 7)
  = 1 – Binom**C**df(n, P, 7)

More Examples:

- more than 4 ~ P(X > 5) = 1 — P(X ≤ 5) = 1 – BinomCdf(n, P, 5)
- no(t) more than 5 ~ P(X ≤ 5) = BinomCdf(n, P, 5)
- greater than 4 ~ P(X > 4) = P(X ≥ 5) = 1 — P(X ≤ 4) = 1 – BinomCdf(n, P, 4)
- no(t) fewer than 5 ~ P(X ≥ 5) = 1 — P(X ≤ 4) = 1 – BinomCdf(n, P, 4)
- less than 5 ~ P(X < 5) = P(X ≤ 4) = BinomCdf(n, P, 4)
- no(t) less than 5 ~ P(X ≥ 5) = 1 — P(X ≤ 4) = 1 – BinomCdf(n, P, 4)
- exceed 5 ~ P(X > 5) = 1 — P(X ≤ 5) = 1 – BinomCdf(n, P, 5)
- at least 5 ~ P(X ≥ 5) = 1 — P(X ≤ 4) = 1 – BinomCdf(n, P, 4)

- at most 5 ~ $P(X \leq 5)$ = BinomCdf(n, P, 5)
- exactly 5 ~ $P(X = 5)$ = BinomPdf(n, P, 5)

For the TI-83 and 84 calculators:

1. Subtract from 1 *only when* the Q relates to > or $\geq$.
2. Use Summation and n*Cx* *only when* the Q relates to $\geq$ or $\leq$.
3. Use numbers for *n* and *x* in *nCx* when the Q relates to =.
4. For Rare–Events, choose the Inequality so that you dont cross the Mean, E(X) = nP.
5. Use BinomPdf for $P(X = a)$ situations.
6. Use BinomCdf for $P(X \leq a)$ situations.

Calculator Clarification! Only when the Q deals with expressions as $P(X > a)$ or $P(X \geq b)$ do we rewrite it as: $1 - P(X \leq c)$.

For expressions as $P(X < a)$ or $P(X \leq b)$, there is ***no need*** to subtract from 1...because the calculator processes $\leq$ probabilities!

Mean and s.d. of a Binomial Distribution [Memorize this!]
Mean: E(X) = nP
Standard Deviation: $\sigma(X) = \sqrt{nPQ}$

Example 1

I Interpret the following situations in Probability Notation and rewrite – if necessary – to make them calculator–ready i.e. describe using *nCx* formulas; **do *not*, however, perform any calculations!**

a) P(at least 3 successes out of 10 trials), P(S) = 0.1
b) P(at most 7 successes out of 12 trials), P(S) = 0.25
c) P(fewer than 6 successes out of 8 trials), P(S) = 0.83
d) P(more than 3 successes out of 9 trials), P(S) = 0.35
e) P(exactly 11 failures out of 15 trials), P(S) = 0.19
f) P(at least 9 failures out of 25 trials), P(S) = 0.75
g) P(not more than 4 failures out of 6 trials), P(S) = 0.42

Solution. If X is a r.v. that denotes # of successes, and Y, the # of failures:

a) $P(X \geq 3) = 1 - P(X \leq 2) = \sum[x = 0 \text{ to } 2]\ 10Cx\,(0.1)^{x}\,(0.9)^{10-x}$

b) $P(X \leq 7) = \sum[x = 0 \text{ to } 7]\ 12Cx\,(0.25)^{x}\,(0.75)^{12-x}$

c) $P(X < 6) = P(X \leq 5) = \sum[x = 0 \text{ to } 5]\ 8Cx\,(0.83)^{x}\,(0.17)^{8-x}$

d) $P(X > 3) = 1 - P(X \leq 3) = 1 - \sum[x = 0 \text{ to } 3]\ 9Cx\,(0.35)^{x}\,(0.65)^{9-x}$

e) Q = 0.81 and $P(Y = 11)$: $15C11\,(0.81)^{11}\,(0.19)^{4}$

or $P(X = 4) = 15C4\,(0.19)^{4}\,(0.81)^{11}$

f) Q = 0.25 and $P(Y \geq 9) = 1 - P(Y \leq 8) = 1 - \sum[y = 0 \text{ to } 8]\ 25Cy\,(0.25)^{y}\,(0.75)^{25-y}$

or $P(X \leq 16) = \sum[x = 0 \text{ to } 16]\ 25Cx\,(0.75)^{x}\,(0.25)^{25-x}$

g) Q = 0.58 and $P(Y \le 4) = 1 - \sum [y = 0 \text{ to } 4]\ 6Cy\ (0.58)^{y}\ (0.42)^{6-y}$

or $P(X \ge 2) = 1 - P(X \le 1) = 1 - \sum [x = 0 \text{ to } 1]\ 6Cx\ (0.42)^{x}\ (0.58)^{6-x}$

Example 2

Suppose we're interested in finding out about the support a candidate John Smith has, and we randomly interview 12 voters. Assume his overall approval rating to be 41%.
a) Describe how the 4 conditions of a Binomial situation are met ***in context***.
b) Define a suitable Binomial r.v., X.
c) What is the probability exactly 6 chaps approve of John Smith?
d) What is the probability more than 8 chaps approve of John Smith?
e) What is the probability fewer than 3 chaps approve of John Smith?
f) What is the probability not more than 5 chaps approve of John Smith?
g) What is the probability of getting at least 1 John Smith supporter?
h) In a sample of 12 individuals, how many would you expect to be John Smith supporters? What is the s.d. of the number of supporters?

i) Suppose voters are repeatedly asked for their preferences. What is the probability that the 1^{st} John Smith voter shall be the 5^{th} one chosen?
j) Suppose voters are asked for their preferences, one after the other. What is the probability that the 4^{th} John Smith voter shall be the 15^{th} one chosen?

Solution.

a)
1. There are 2 outcomes: an individual is a John Smith supporter or not;
2. There are a fixed number of "trials", n = 12;
3. The probability of success i.e. being a John Smith supporter, is constant, P = 0.41;
4. The outcomes [supporter / The Devil] are independent of each other since there are at least 120 individuals in the population $[N \ge 10n = 10 \cdot 12]$

b) X is a r.v. denoting Number of John Smith supporters in a sample of 12; $X \sim B(12, 0.41)$

c) $P(X = 6) = {}^{12}C6\ (0.41)^{6}(0.59)^{6} = 18.51\%$ [**Use the BinomPdf command because of the EQUAL TO sign…]**

d) $P(X > 8)$
$= 1 - P(X \le 8)$ **Use a Number Line to see why!**
$= 1 - (x = 0)\sum(x = 8)\ 12Cx\ (0.41)^{x}\ (0.59)^{(12-x)}$ [**Use the BinomCdf command because of the LESS THAN OR EQUAL TO sign…]**
= 1.82%

e) $P(X < 3)$
$= P(X \le 2)$ **Use a Number Line to see why!**
$= (x = 0)\sum(x = 2)\ 12Cx\ (0.41)^{x}\ (0.59)^{(12-x)}$ [**Use the BinomCdf command because of the LESS THAN OR EQUAL TO sign…]**
= 7.34%

f) $P(X \le 5)$
$= (x = 0)\sum(x = 5)\ 12Cx\ (0.41)^{x}\ (0.59)^{(12-x)}$ [**Use the BinomCdf command because of the LESS THAN OR EQUAL TO sign…]**

= 63.84%
g) $P(X \geq 1)$
= 1 – $P(X = 0)$ **Use a Number Line to see why!**
= 1 – $^{12}C0\ (0.41)^{0}(0.59)^{12}$ **[Use the BinomCdf command because of the LESS THAN OR EQUAL TO sign…]**
= 99.82%
h) $E(X) = nP = 12(0.41) = 4.92$
$\sigma(X) = \sqrt{nPQ} = \sqrt{12(0.41)(0.59)} = 1.7038$
i) This is not a Binomial event since the number of trials is not fixed…this is just a simple Probability Rules situation. If A ~ event that a voter is a John Smith supporter, required: P(A', A', A', A', A) = $P(A')^{4}$ P(A), assuming independence of voting preferences = $(1 – 0.41)^{4}$ (0.41) = 4.96%

j) If the 4th John Smith voter shall be the 15th one chosen, then we must have 3 John Smith voters amongst the 1st 14…and the last (15th) fellow must be a John Smith voter!
The 1st of these is a Binomial event: X is a r.v. denoting Number of John Smith supporters in a sample of 14; X ~ B(14, 0.41)...and the 2nd probability is simply 0.41!

Required: $P(X = 3)\times 0.41$ **[Use the BinomPdf command because of the EQUAL TO sign…]**
= **[**$^{14}C3\ (0.41)^{3}(0.59^{12})$**]×0.41**
= 3.1%

Example 3

Suppose 26% of drivers in a city are driving without seat–belts. If we randomly select 15 drivers on a weekend,
a) In phrases [only] and in context, describe how the 4 conditions of a Binomial situation are met ***in context***.
b) Define a suitable Binomial r.v., X.
c) What is the probability exactly 3 drivers are driving without seat–belts?
d) What is the probability more than 6 drivers are driving without seat–belts?
e) What is the probability fewer than 5 drivers would be driving without seatbelts?
f) What is the probability not more than 5 drivers are driving without seat–belts?
g) What is the probability of getting at least 2 drivers that are driving without seat–belts?
h) How many drivers would you expect to be driving without seat–belts? What is the s.d. of the number of drivers?
i) Suppose drivers are observed for their seat–belt compliance. What is the probability that the 1st driver not wearing his seat–belt is is the 4th one chosen?
j) Suppose drivers are observed for their seat–belt compliance, one after the other. What is the probability that the 6th driver not wearing a seat–belt is the 10th one chosen?
Suppose 26% of drivers in a city are driving without seat–belts. If we randomly select 15 drivers on a weekend,

Solution.

a)

- 2 outcomes: driver wearing seatbelt / not;
- P(wearing seatbelt) = 0.26 = constant;
- incidence of wearing seat–belt is independent of the drivers
- fixed number of trials, n = 12

b) X ~ r.v. denoting # of drivers not wearing a seat–belt amongst 15: X ~ Bin(15, 0.26)

c) P(X = 3) = 21.56% **Show nCx notation**.

d) $P(X > 6) = 1 - P(X \le 6)$ **Use a Number Line to see why...** = 6.83% **Show nCx notation**.

e) $P(X < 5) = P(X \le 4)$ **Use a Number Line to see why...** = 65.31% **Show nCx notation**.

f) $P(X \le 5)$ = 82.87% **Show nCx notation**.

g) $P(X \ge 2) = 1 - P(X \le 1)$ **Use a Number Line to see why...** = 93.14% **Show nCx notation**.

h) E(X) = nP = 3.9, σ(X) = √nPQ = √15(0.26)(0.74) = 1.6988 **Show nCx notation**.

i) This is not a Binomial event since the number of trials is not fixed…this is just a simple Probability Rules situation. If A ~ event that a driver is not wearing a seat–belt, required: P(A', A', A', A) = $P(A')^3$ P(A), assuming independence of seat–belt wearing behaviour

= $(1 - 0.26)^3$ (0.26) = 10.53%

j) If the 6th driver not wearing a seat–belt shall be the 10th one chosen, then we must have 5 non–seat–belt drivers amongst the 1st 9!…and the last (10th) fellow must be a non–seat–belt driver! The 1st of these is a Binomial event: X is a r.v. denoting Number of drivers not wearing seat–belts in a sample of 9; X ~ B(9, 0.26)...and the 2nd probability is simply 0.26!

Required: P(X = 5)×0.26 [**Use the BinomPdf command because of the EQUAL TO sign…]**

= **[** $^{9}C5\ (0.26)^{5}(0.74)^{4}$ **]×0.26**

= 1.16%

Example 4

Suppose we're interested in finding out about the support a candidate John Smith has, and we randomly interview 12 individuals. Assume his overall approval rating to be 41%.

a) What is the probability that at least 7 chaps approve of John Smith?

b) What is the probability that more than 6 chaps ***dis***approve of John Smith?

c) What is the probability that fewer than 3 chaps ***dis***approve of John Smith?

Solution.

a) Let X is a r.v. denoting Number of Obama supporters in a sample of 12;

X ~ B(12, 0.41)

P(X > 7)

= $1 - P(X \le 6)$

= $1 - (x = 0)\sum(x = 6)\ 12Cx\ (0.41)^{x}\ (0.59)^{(12 - x)}$

= 17.64%

b) **Method I** [a little complicated!]

More than 6 chaps ***dis***approving of John Smith

~ 7, 8...11, 12 chaps ***dis***approving
~ 5, 4, ...1, 0 chaps approving of John Smith [because we have 12 chaps in all!].
$P(X \le 5)$
$= (x = 0)\sum(x = 5)\ 12Cx\ (0.41)^{x}\ (0.59)^{(12 - x)}$
= 63.84%
Method II [very elegant!!]
Let Y be a r.v. denoting # of individuals who ***dis***approve of John Smith
Y ~ B(12, **0.59**)
Required: $P(Y > 6)$
$= 1 - P(Y \le 6)$
$= 1 - (y = 0)\sum(y = 6)\ 12Cy\ (0.59)^{y}\ (0.41)^{(12 - y)}$
= 63.84%%
c) $P(Y < 3)$
$= P(Y \le 2)$
$= (y = 0)\sum(y = 2)\ 12Cy\ (0.59)^{y}\ (0.41)^{(12 - y)}$
= 0.35%

Example 5

Historically, the pass rate for AP Calculus AB has been about 59.1%. Suppose we randomly select 150 AP Calculus AB students.
a) Describe how the 4 conditions of a Binomial situation are met ***in context***.
b) Define a suitable Binomial r.v., X.
c) What is the probability exactly 94 students pass the AB Exam?
d) What is the probability more than 100 students pass the AB Exam?
e) What is the probability fewer than 80 students pass the AB Exam?
f) What is the probability at least 35 students ***fail*** the AB Exam?
g) What is the probability that between 86 and 110 students pass the AB Exam?
h) How many would you expect to pass the AB Exam? What is the s.d. of the number of
i) Suppose students are repeatedly asked for their performance [pass / fail the Calculus AB Exam]. What is the probability that the 1^{st} student that passed the Exam is the 6^{th} one chosen?
j) Suppose students are repeatedly asked for their performance [pass / fail the Calculus AB Exam]. What is the probability that the 10^{th} student that passed is the 17^{th} one chosen?

Solution.

a)

- 2 outcomes: student passes / fails the AB Exam;
- P(passing AB Exam) = 0.591 = constant;
- Incidence of passing the Exam is independent of each other since there are at least 1500 students [$N > 10n = 10 \cdot 150$] taking the AB Exam
- fixed number of trials, n = 150 students

b) X ~ r.v. denoting # of Calculus AB students amongst 150: X ~ Bin(150, 0.591)
c) $P(X = 94) = 150C94\ (0.591)^{94}\ (0.90)^{(56)}$ = **Finish**.
d) $P(X > 100) = 1 - P(X \le 100) = 1 - (x = 0)\sum(x = 100)\ 150Cx\ (0.591)^{x}\ (1 - 0.591)^{(150 - x)}$ = **Finish.**

e) $P(X < 80) = P(X \leq 79) = (x = 0)\sum(x = 79)\ 150Cx\ (0.591)^{x}\ (1 - 0.591)^{(150 - x)}$ **Finish.**
f) If Y is r.v. Denoting the Number of who that Fail the AB Exam, $Y \sim Bin(150, 1 - 0.409)$.

Required, $P = P(Y \geq 35) = 1 - P(Y \leq 34) = (y = 0)\sum(y = 34)\ 150Cy\ (0.409)^{y}\ (0.591)^{(150 - y)}$.

g) $P(86 \leq X \leq 110) = (x = 0)\sum(x = 110)\ 150Cx\ (0.591)^{x}\ (1 - 0.591)^{(150 - x)} - [(x = 0)\sum(x = 85)\ 150Cx\ (0.591)^{x}\ (1 - 0.591)^{(150 - x)}]$
h) $E(X) = nP = 150(0.591) =$ **Finish;** $\sigma(X) = \sqrt{nPQ} = \sqrt{150(0.591)(1 - 0.591)} =$ **Finish.**

i)If $A \sim$ event that a student passes the AB Exam, required: $P(A', A', A', A', A', A) = P(A')^{5}\ P(A)$, assuming independence of scores

$= (1 - 0.591)^{5}\ (0.591) =$ **Finish.**

j) If the 10th student that passed is the 17th one chosen, then we <u>must</u> have 9 students that passed amongst the 1st 16 students!...and the last (17th) fellow <u>must</u> have passed, too!
The 1st of these is a Binomial event: $X \sim B(16, 0.591)$...and the 2nd probability is simply 0.591!

Required, $P = \mathbf{[}{}^{\mathbf{16}}C9\ (0.591)^{9}(0.409)^{7}\mathbf{]\times 0.591}$
= **Finish.**