Chapter 6

Discrete Probability Distributions

Distribution, mean and standard deviation of discrete random variables are described, first in general, then for the binomial and Poisson special cases.

6.1 Discrete Random Variables

A *random variable*, denoted by a capital letter such as X, is a "rule" which assigns a number to each outcome in sample space of a probability experiment[1]. A random variable is *discrete* if outcomes assigned to finite or countably infinite real values.

Exercise 6.1 (Discrete Random Variables)

1. *Properties of discrete probability distribution of random variable X, value x.*
   - $\sum P(x) = 1, \quad 0 \leq P(x) \leq 1$
   - *expected value, mean*: $\mu_X = \sum [x \cdot P(x)]$
   - *variance*: $\sigma_X^2 = \sum [(x-\mu)^2 P(x)] = \sum [x^2 \cdot P(x)] - \mu_X^2$
   - *standard deviation*: $\sigma_X = \sqrt{\sigma_X^2}$

2. *Discrete or Continuous?*

   (a) **discrete** / **continuous**. Number of seizures in a year.

   (b) **discrete** / **continuous**. Waiting time at Burger King.

   (c) **discrete** / **continuous**. Temperature in Michigan City.

   (d) **discrete** / **continuous**. Number of bikes on bike rack.

   (e) **discrete** / **continuous**. Number of heads in three coin tosses.

   (f) **discrete** / **continuous**. Number of pips in roll of dice.

[1] More technically, a random variable is a *function* mapping from sample space to real line.

(g) **discrete** / **continuous**. Patient's ages.

3. *Sample distribution versus probability distribution: seizures.*

(a) *Sample distribution.*
If data from 100 epileptic people sampled at random in one year was

number seizures	number people
0	17
2	21
4	18
6	11
8	16
10	17

observed *average* number of seizures would be

$$\begin{aligned}\bar{x} &= \frac{17(0)+21(2)+18(4)+11(6)+16(8)+17(10)}{100} \\ &= 0\cdot\frac{17}{100}+2\cdot\frac{21}{100}+4\cdot\frac{18}{100}+6\cdot\frac{11}{100}+8\cdot\frac{16}{100}+10\cdot\frac{17}{100} \\ &= 0(0.17)+2(0.21)+4(0.18)+6(0.11)+8(0.16)+10(0.17) \ =\end{aligned}$$

which is equal to (circle one) **4.32** / **4.78** / **5.50** / **5.75**.

(StatCrunch: Relabel var1 as seizures, var2 as number. Type data into seizures and number columns. Stat, Summary Stats, Grouped/Binned data, Bins in: seizures, Counts in: number, Statistics: Mean, Std. dev., Compute. Notice sample $\bar{x} = 4.78$.)
Sample average $\bar{x}$ is a (circle one) **parameter** / **statistic**.

Observed standard deviation in number of seizures,

$$\begin{aligned}s &= \sqrt{\frac{17(0-4.78)^2+\cdots+17(10-4.78)^2}{100-1}} \\ &= \sqrt{(0-4.78)^2\cdot\frac{17}{100-1}+\cdots+(10-4.78)^2\cdot\frac{17}{100-1}} \ \approx\end{aligned}$$

(circle one) **3.32** / **3.49** / **3.50** / **3.75**,
(StatCrunch: Notice sample $s \approx 3.49$.)
Sample SD s is also a (circle one) **parameter** / **statistic**.

(b) *Probability distribution.*
Table of probability distribution:

number seizures, x	$P(x)$
0	$\frac{17}{100} = 0.17$
2	0.21
4	0.18
6	0.11
8	0.16
10	0.17

(Blank data table. Relabel var1 as seizures, var2 as P(x). Type data into seizures and P(x) columns. Data, Save data, 6.1.2. seizures probability distribution. Stat, Calculators, Custom, Values in: seizures, Weights in: P(x), Okay. Notice population $\sigma_x \approx 3.47$ given here $\neq$ previous sample $s_x \approx 3.49$.)

Probability distribution for number of seizures, X, is (circle one)

i. *only* associated with random sample of one hundred epileptic people.

ii. a *guess (hypothesis)* of the chances associated with different number of seizures in the long run.

This probability distribution is what we (circle one) **expect** / **observe**. Since probability distribution is associated with population, both mean, $\mu = 4.78$, and standard deviation, $\sigma = 3.47$, are **parameters** / **statistics**.

Probability histogram (graph) of distribution.

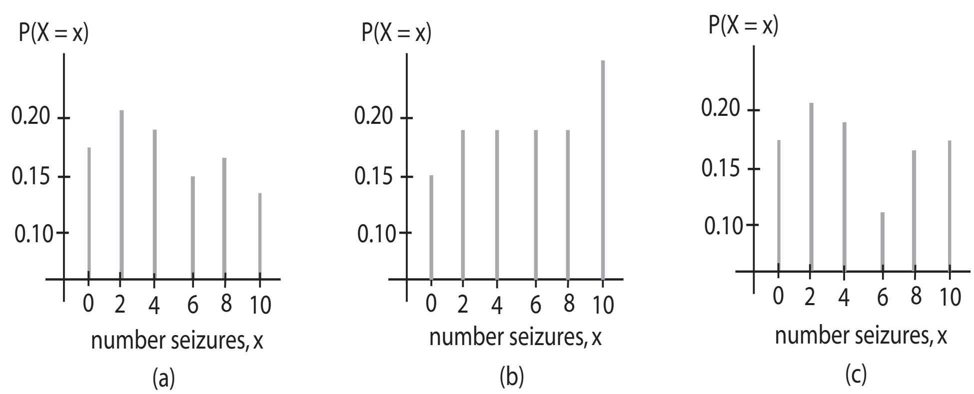

Figure 6.1 (Probability histogram: seizures)

Which of three probability histograms describes probability distribution of number of seizures? Choose one. **(a)** / **(b)** / **(c)**

Probability Function, number of seizures. Choose one.

i. Function (a).

$$P(X = x) = \begin{cases} 0.17, & \text{if } x = 0 \\ 0.21, & \text{if } x = 2 \end{cases}$$

ii. Function (b).

$$P(X = x) = \begin{cases} 0.18, & \text{if } x = 4 \\ 0.11, & \text{if } x = 6 \end{cases}$$

iii. Function (c).

$$P(X = x) = \begin{cases} 0.17, & \text{if } x = 0 \\ 0.21, & \text{if } x = 2 \\ 0.18, & \text{if } x = 4 \\ 0.11, & \text{if } x = 6 \\ 0.16, & \text{if } x = 8 \\ 0.17, & \text{if } x = 10 \end{cases}$$

4. *Probability distribution, mean μ_X and SD σ_X: number of seizures, X*

number seizures, x	$P(x)$
0	0.17
2	0.21
4	0.18
6	0.11
8	0.16
10	0.17

(a) *Various probabilities associated with number of seizures.*

i. Chance a person has 8 epileptic seizures is
$P(8) = P(X = 8) =$ (circle one) **0.14** / **0.15** / **0.16** / **0.17**.

ii. Chance a person has *at most* 4 seizures is
$P(X \leq 4) = P(0) + P(2) + P(4) =$
(circle one) **0.17** / **0.21** / **0.56** / **0.67**.

iii. Chance a person has *at least* 4 seizures is
$P(X \geq 4) = P(4) + P(6) + P(8) + P(10) = 1 - P(X \leq 3) =$
(circle one) **0.21** / **0.38** / **0.56** / **0.62**.

iv. $P(0) + P(2) + P(4) + P(6) + P(8) + P(10) =$ **0.97** / **0.98** / **1**.

v. $P(2.1) =$ (circle one) **0** / **0.21** / **0.56** / **0.67**.

(b) *Mean (expected value) of number of seizures.*

$$\begin{aligned} \mu_X &= \sum [x \cdot P(x)] \\ &= 0 \cdot P(0) + 2 \cdot P(2) + 4 \cdot P(4) + 6 \cdot P(6) + 8 \cdot P(8) + 10 \cdot P(10) \\ &= 0(0.17) + 2(0.21) + 4(0.18) + 6(0.11) + 8(0.16) + 10(0.17) \ = \end{aligned}$$

(circle one) **4.32** / **4.78** / **5.50** / **5.75**.

(Stat, Calculators, Custom, Values in: seizures, Weights in: P(x), Okay. Notice μ_X = Mean: 4.78.)

Understanding mean (expected value): point of balance.

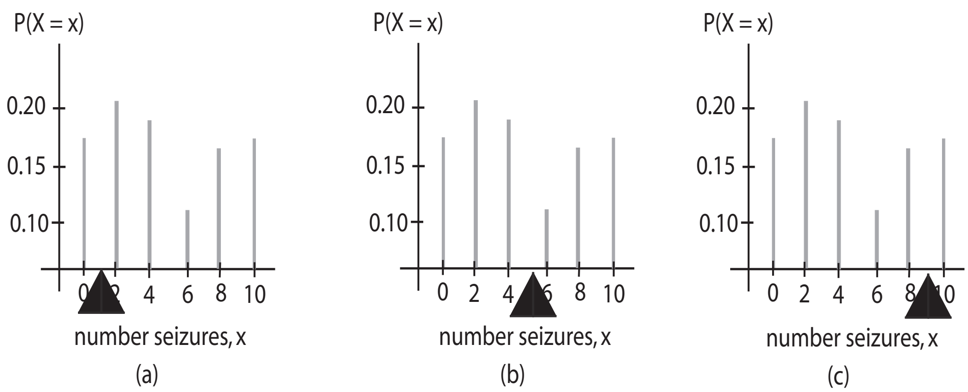

Figure 6.2 (Mean number of seizures: point of balance.)

Mean balances "weight" of probability in graph **(a)** / **(b)** / **(c)**.
In other words, mean (expected value) close to (circle one) **1** / **5** / **9**.

(c) *Standard deviation in number of seizures.*

$$\sigma_X = \sqrt{\sum [(x - \mu_X)^2 P(x)]}$$
$$= \sqrt{(0-4.78)^2(0.17) + (2-4.78)^2(0.21) + \cdots + (10-4.78)^2(0.17)} \approx$$

(circle one) **3.47** / **4.11** / **5.07** / **6.25**.

(Stat, Calculators, Custom, Values in: seizures, Weights in: P(x), Okay. Notice σ_X = Std, Dev.: 3.47.)

Understanding standard deviation: dispersion (spread).

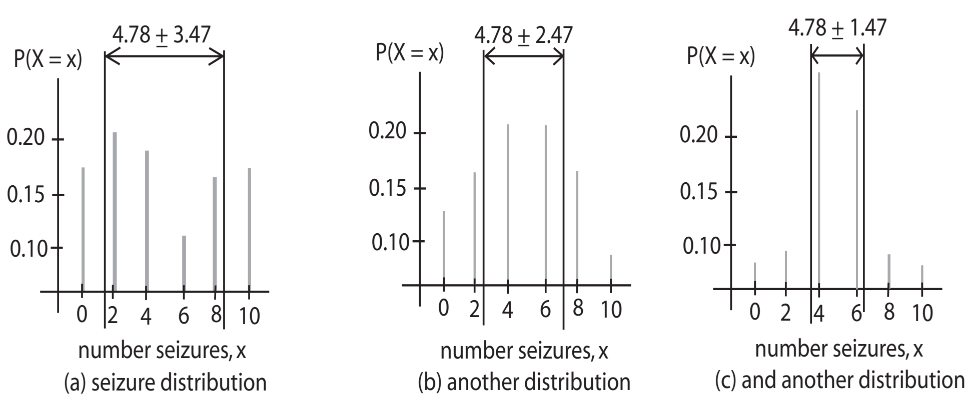

Figure 6.3 (SD in number of seizures: dispersion (spread).)

Standard deviation measures dispersion of a probability distribution.
Most dispersed distribution occurs in **(a)** / **(b)** / **(c)**. We expect to see about 4.78 "±" 3.47 seizures according to seizure probability distribution.

Variance in number of seizures is
$\sigma^2 \approx 3.47^2 \approx$ (circle one) **10.02** / **11.11** / **12.07** / **13.25**.

(Stat, Calculators, Custom, Values in: seizures, Weights in: P(x), *choose Variance*, Okay.)

5. *Probability distribution, mean μ_X and SD σ_X: number of bikes on bike rack, X*

number bikes, x	$P(x)$
5	$\frac{1}{5} = 0.2$
6	0.2
7	0.2
8	0.2
9	0.2

For example, there is a 20% chance 6 bikes are on bike rack.

(a) *Probability histogram, number of bikes.*

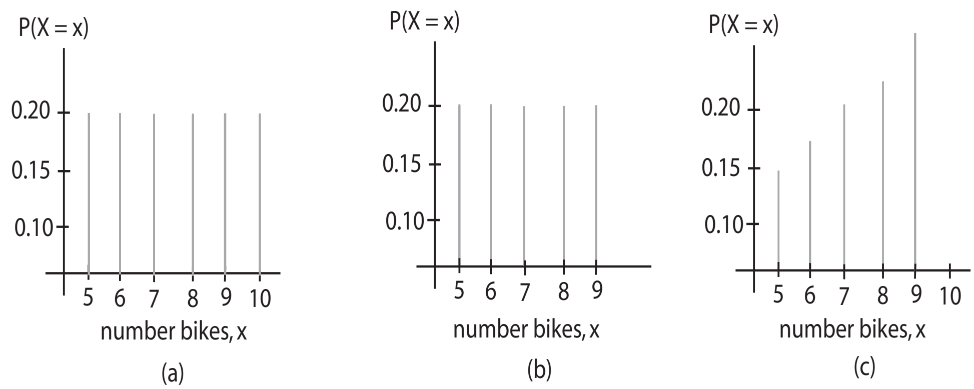

Figure 6.4 (Probability histogram of number of bikes.)

Probability histogram, number of bikes: (choose one) **(a)** / **(b)** / **(c)**.

(b) *Probability function, number of bikes.* Choose *two*!

i. Function (a).

$$P(X = x) = \begin{cases} 0.2, & \text{if } x = 5, \\ 0.2, & \text{if } x = 6, \\ 0.2, & \text{if } x = 7, \\ 0.2, & \text{if } x = 8, \\ 0.2, & \text{if } x = 9. \end{cases}$$

ii. Function (b).

$$P(X = x) = \frac{1}{5} = 0.2, \quad x = 5, 6, 7, 8, 9.$$

iii. Function (c).

$$P(X = x) = \frac{x}{35}, \quad x = 5, 6, 7, 8, 9.$$

(c) *Various probabilities associated with number of bikes.*

i. Chance bike rack has 8 bicycles is
$P(8) =$ (circle one) $\mathbf{\frac{1}{5}}$ / $\mathbf{\frac{2}{5}}$ / $\mathbf{\frac{3}{5}}$ / $\mathbf{\frac{4}{5}}$.

ii. Chance bike rack has *at most* 6 bicycles is
$P(X \leq 6) = P(5) + P(6) =$ (circle one) $\mathbf{\frac{1}{5}}$ / $\mathbf{\frac{2}{5}}$ / $\mathbf{\frac{3}{5}}$ / $\mathbf{\frac{4}{5}}$.

iii. Chance bike rack has *at least* 6 bicycles is
$P(X \geq 6) = P(6) + P(7) + P(8) + P(9) = 1 - P(X \leq 5) =$
(circle one) $\mathbf{\frac{1}{5}}$ / $\mathbf{\frac{2}{5}}$ / $\mathbf{\frac{3}{5}}$ / $\mathbf{\frac{4}{5}}$.

iv. Chance bike rack has *more than* 6 bicycles is
$P(X > 6) = P(X \geq 7) = P(7) + P(8) + P(9) = 1 - P(X \leq 6) =$
(circle one) $\mathbf{\frac{1}{5}}$ / $\mathbf{\frac{2}{5}}$ / $\mathbf{\frac{3}{5}}$ / $\mathbf{\frac{4}{5}}$.

(d) *Mean (expected value) of number of bikes.*

$$\begin{aligned} \mu_X &= \sum [x \cdot P(x)] \\ &= 5\times\frac{1}{5} + 6\times\frac{1}{5} + 7\times\frac{1}{5} + 8\times\frac{1}{5} + 9\times\frac{1}{5} = \end{aligned}$$

(circle one) **5** / **6** / **7** / **8**.

(StatCrunch: Blank data table. Relabel var1 bikes, var2 P(x). Data, Data save 6.1.3 bike distribution. Stat, Calculators, Custom, Values in: bikes, Weights in: P(x), Okay. Notice Mean: 7 or notice 7 is balance point of probability histogram.)

(e) *Standard deviation in number of bikes.*

$$\begin{aligned} \sigma_X &= \sqrt{\sum [(x - \mu_X)^2 P(x)]} \\ &= \sqrt{(5-7)^2(0.2) + (6-7)^2(0.2) + \cdots + (9-7)^2(0.2)} \approx \end{aligned}$$

(circle one) **1.41** / **2.41** / **3.07** / **4.25**.

(StatCrunch: Notice Std. Dev.: 1.41.)

In other words, we expect to see about 7 "±" 1.4 bikes on bike rack.

6. *Probability distribution, mean μ_X and SD σ_X: roulette payoff, X*
Roulette table has 38 numbers: numbers are 1 to 36, 0 and 00. A ball is spun on a roulette wheel. After a time, ball drops into one of 38 slots which correspond to 38 numbers on roulette table.

(a) *Betting on even.* Let random variable X be payoff from \$1 bet on even: \$1 lost if ball drops on odd or 0 or 00, \$1 won (added to \$1 bet) if even.

payoff, x	$P(x)$
-\$1	$\frac{20}{38}$
\$1	$\frac{18}{38}$

Mean is $\mu_X = -1 \times \frac{20}{38} + 1 \times \frac{18}{38} = -\frac{2}{38}$ (≈ -0.05) and so

$$\sigma_X = \sqrt{\left(-1 - \left(-\frac{2}{38}\right)\right)^2 \frac{20}{38} + \left(1 - \left(-\frac{2}{38}\right)\right)^2 \frac{18}{38}} \approx$$

(circle one) **0.051** / **0.999** / **1.573** / **2.251**

(StatCrunch: Blank data table. Relabel var1 payoff, var2 frequency, var3 P(x). Since StatCrunch deals with fractions awkwardly, first type -1, 1 into payoff, and 20, 18 into frequency, then, to derive probabilities, Data, Compute expression, Expression: frequency/38, New column name: P(x), Compute. Data, Data save 6.1.4 roulette even distribution. Stat, Calculators, Custom, Values in: payoff, Weights in: P(x), Okay. Notice Std. Dev.: 0.999.)

We expect to lose a nickel "±" a dollar betting \$1 on even.

Also, $\sigma_X^2 \approx 0.999^2 \approx$ (circle one) **0.997** / **0.998** / **0.999** / **1.000**.

(Stat, Calculators, Custom, Values in: payoff, Weights in: P(x), *choose Variance*, Okay.)

(b) *Betting on a section.* Let random variable Y be payoff from a \$1 bet on a section (with 12 numbers): \$1 lost if ball drops on one of 24 numbers not in section or 0 or 00, \$2 won (added to \$1 bet) if number in section.

payoff, x	$P(x)$
-\$1	$\frac{26}{38}$
\$2	$\frac{12}{38}$

Mean is $\mu_Y = -1 \times \frac{26}{38} + 2 \times \frac{12}{38} = -\frac{2}{38}$ (≈ -0.05) and so

$$\sigma_Y = \sqrt{\left(-1 - \left(-\frac{2}{38}\right)\right)^2 \frac{26}{38} + \left(2 - \left(-\frac{2}{38}\right)\right)^2 \frac{12}{38}} \approx$$

(circle one) **0.05** / **0.47** / **1.39** / **2.25**

(StatCrunch: Blank data table. Relabel var1 payoff, var2 frequency, var3 P(x). Since StatCrunch deals with fractions awkwardly, first type -1, 2 into payoff, and 26, 12 into frequency, then, to derive probabilities, Data, Compute expression, Expression: frequency/38, New column name: P(x), Compute. Data, Data save 6.1.5 roulette section distribution. Stat, Calculators, Custom, Values in: payoff, Weights in: P(x), Okay. Notice Std. Dev.: 1.39.)

We expect to lose a nickel "±" \$1.39 betting \$1 on a section.

Also, $\sigma_Y^2 \approx 1.39^2 =$ (circle one) **1.945** / **1.946** / **1.947** / **1.948**.

What is expected on average when betting \$100 on a section?
$\mu_Y = -100 \times \frac{26}{38} + 200 \times \frac{12}{38} \approx$ (circle one) **−5.26** / **−9.99** / **−15.73**

6.2 The Binomial Probability Distribution

Probability function for *binomial* is

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

with expected value, variance and standard deviation,

$$\mu_X = np; \quad \sigma_X^2 = np(1-p); \quad \sigma_X = \sqrt{np(1-p)}.$$

Probability distribution is binomial if

- n trials, where n is fixed in advance,
- trials have two possible outcomes: success or failure,
- trials independent of one another,
- probability of success same for each trial.

Exercise 6.2 (The Binomial Probability Distribution)

1. *Binomial probability distribution, mean μ_X, SD σ_X: number of cases won, X*
   A lawyer estimates she wins 40% ($p = 0.4$) of her cases. Assume each trial is independent of one another and, in general, this problem obeys conditions of a binomial experiment. The lawyer is currently involved in 10 ($n = 10$) cases. For example, there is a 4.3% chance she wins 7 of her 10 cases.

number cases won, x	$P(x)$
0	0.006
1	0.040
2	0.121
3	0.215
4	0.251
5	0.201
6	0.111
7	0.043
8	0.011
9	0.002
10	0.000

   (a) *Probability histogram, number of wins.*

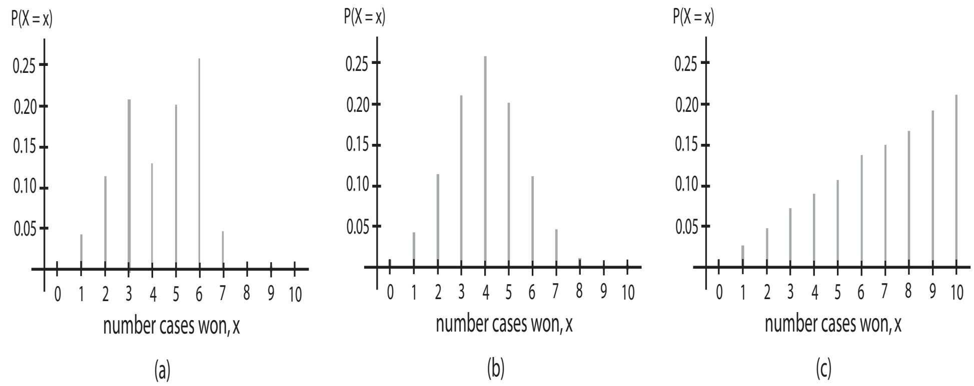

Figure 6.5 (Probability histogram of number of wins.)

Probability histogram, number of wins: (choose one) **(a)** / **(b)** / **(c)**.

(b) *Probability function, number of wins.* Choose *two!*

i. Function (a).

$$P(X = x) = \frac{x}{46}, \quad x = 0, 1, 2, \ldots, 10.$$

(Type 0,1,2 …, 10 into L_1. Define $L_2 = L_1 \div 46$ ENTER. Probabilities in L_2 equal to $P(x)$?)

ii. Function (b).

$$P(X = x) = \begin{cases} 0.006, & \text{if } x = 0, \\ 0.040, & \text{if } x = 1, \\ 0.121, & \text{if } x = 2, \\ \vdots & \vdots \\ 0.002, & \text{if } x = 9, \\ 0.000, & \text{if } x = 10. \end{cases}$$

iii. Function (c).

$$P(X = x) = {}_{10}C_x 0.4^x (1 - 0.4)^{10-x}, \ x = 0, 1, 2 \ldots, n$$

(StatCrunch: Blank data table. Relabel var1 case, var2 P(x). Type number of cases, 0, 1, ..., 10 into case column. To derive probabilities, Data, Compute expression, Expression: dbinom(case,10,0.4), New column name: P(x), Compute. Data, Data save 6.2.1 binomial lawyer distribution.)

(c) *Various probabilities associated with number of wins.*

i. Chance lawyer wins 8 cases is

$$P(8) = {}_{10}C_8 p^8 (1-p)^{10-8} = \frac{10!}{8!(10-8)!}(0.4)^8(1-0.4)^2 = 45 \cdot 0.4^8 0.6^2 \approx$$

(circle one) **0.006** / **0.011** / **0.040** / **0.121**.

(StatCrunch: Stat, Calculators, Binomial, n: 10, p: 0.4, Prob(X = 8) (not Prob(X $\leq$ 8)!) Compute. Notice Prob(X = 8) = 0.011.)

ii. Chance lawyer has *at most* 6 wins is

$$\begin{aligned} P(X \leq 6) &= P(0) + P(1) + P(2) + P(3) + P(4) + P(5) + P(6) \\ &= {}_{10}C_0 p^0 (1-p)^{10-0} + {}_{10}C_1 p^1 (1-p)^{10-1} + \cdots + {}_{10}C_6 p^6 (1-p)^{10-6} \\ &= \frac{10!}{0!(10-0)!}(0.4)^0(1-0.4)^{10} + \cdots + \frac{10!}{6!(10-6)!}(0.4)^6(1-0.4)^4 \\ &= 1 \cdot 0.4^0 0.6^{10} + 10 \cdot 0.4^1 0.6^9 + \cdots + 210 \cdot 0.4^6 0.6^4 \approx \end{aligned}$$

(circle one) **0.834** / **0.934** / **0.945** / **0.993**.

(Long way: 1 × 0.4^(0) × 0.6^(10) + 10 × 0.4^(1) × 0.6^(9) + · · · + 210 × 0.4^(6) × 0.6^(4) OR

Short way: Stat, Calculators, Binomial, n: 10, p: 0.4, Prob(X ≤ 6) Compute.)

iii. Chance lawyer has *less than* 6 wins is
$P(X < 6) = P(X \leq 5) \approx$ (circle one) **0.834** / **0.934** / **0.945** / **0.993**.

(StatCrunch: Stat, Calculators, Binomial, n: 10, p: 0.4, Prob(X < 6) Compute.)

(d) *Mean (expected value) of number of wins.*
long way:

$$\begin{aligned} \mu_X &= \sum [x \cdot P(x)] \\ &= 0 \times 0.006 + 1 \times 0.040 + \cdots + 10 \times 0.000 \approx \end{aligned}$$

(circle one) **3** / **4** / **5** / **6**.

(StatCrunch: Stat, Calculators, Custom, Values in: case, Weights in: P(x), Compute. Notice, Mean: 4)

short way:

$$\mu_X = np = 10 \times 0.4 =$$

(circle one) **3** / **4** / **5** / **6**.

(e) *Standard deviation in number of wins.*
long way:

$$\begin{aligned} \sigma_X &= \sqrt{\sum [(x - \mu_X)^2 P(x)]} \\ &= \sqrt{(0-4)^2(0.006) + (1-4)^2(0.040) + \cdots + (10-4)^2(0.000)} \approx \end{aligned}$$

(circle one) **1.23** / **1.44** / **1.55** / **1.76**.

(Stat, Calculators, Custom, Values in: case, Weights in: P(x), Compute. Notice, Std. Dev.: 1.59)

short way:

$$\sigma_X = \sqrt{np(1-p)} = \sqrt{10(0.4)(1-0.4)} \approx$$

(circle one) **1.23** / **1.44** / **1.55** / **1.76**.

We expect lawyer to win about 4 "±" 1.6 of 10 trials.

(f) *Binomial experiment?* Match columns.

binomial conditions	lawyer example
(a) n trials, n is fixed in advance of experiment.	(A) $p = 0.4$ chance lawyer wins each trial.
(b) Trials have possible outcomes: success or failure.	(B) Each trial is independent of one another.
(c) Trials are independent of one another.	(C) There are $n = 10$ trials.
(d) Probability of success same for each trial.	(D) Trials can only be won or lost.

binomial conditions	(a)	(b)	(c)	(D)
lawyer example	(C)			

2. *Binomial: number of airplane engine failures.*
Each engine of four ($n = 4$) on an airplane fails 11% ($p = 0.11$) of the time. Assume this problem obeys conditions of a binomial experiment.

(a) Chance two engines fail is
$P(2) = {}_4C_2 \times (0.11)^2 \times (1-0.11)^{4-2} \approx$ (circle one) **0.005** / **0.011** / **0.058**.
(StatCrunch: Stat, Calculators, Binomial, n: 4, p: 0.11, Prob(X = 2) (not Prob(X ≤ 2)!) Compute.)

(b) Chance three engines fail is
$P(3) = {}_4C_3 \times (0.11)^3 \times (1-0.11)^{4-3} \approx$ (circle one) **0.005** / **0.011** / **0.040**.
(StatCrunch: Stat, Calculators, Binomial, n: 4, p: 0.11, Prob(X = 3) Compute.)

(c) Chance *at most* two engines fail is
$P(X \leq 2) = P(0) + P(1) + P(2) \approx$ (circle one) **0.995** / **0.997** / **0.999**.
(StatCrunch: Stat, Calculators, Binomial, n: 4, p: 0.11, Prob(X ≤ 2) Compute.)

(d) Chance *less than* two engines fail is (circle one)
$P(X < 2) = P(X \leq 1) = P(0) + P(1) \approx$ **0.938** / **0.997** / **0.999**.
(StatCrunch: Stat, Calculators, Binomial, n: 4, p: 0.11, Prob(X < 2) Compute.)

(e) *Expected number of failures*
$\mu_X = np = 4(0.11) =$ **0.44** / **0.51** / **0.62**.

(f) *Standard deviation in number of failures*
$\sigma_X = \sqrt{4(0.11)(1-0.11)} \approx$ **0.45** / **0.56** / **0.63**.
So, expect 0.44 "±" 0.63 failures.

3. *Binomial: number of widget defects.*
Each of fourteen randomly chosen widgets are defective 21% of the time.

(a) *Binomial experiment?* Match columns.

binomial conditions	widget example
(a) n trials, n is fixed in advance of experiment.	(A) $p = 0.21$ chance widget is defective.
(b) Trials have possible outcomes: success or failure.	(B) $n = 14$ widgets chosen.
(c) Trials are independent of one another.	(C) each widget is defective or not.
(d) Probability of success same for each trial.	(D) each widget chosen independent of another.

binomial conditions	(a)	(b)	(c)	(D)
widget example				

(b) Chance seven widgets defective
$P(7) = {}_{14}C_7 \times (0.21)^7 \times (1 - 0.21)^{14-7} \approx$ **0.005 / 0.012 / 0.040**.
(StatCrunch: Stat, Calculators, Binomial, n: 14, p: 0.21, Prob(X = 7) Compute.)

(c) Chance *at most* ten widgets defective
$P(X \leq 10) = P(0) + P(1) + \cdots + P(10) \approx$ **0.995 / 0.997 / 0.999**.
(StatCrunch: Stat, Calculators, Binomial, n: 14, p: 0.21, Prob(X $\leq$ 10) Compute.)

(d) Chance *at least* ten widgets defective
$P(X \geq 10) = P(10) + P(11) + P(12) + P(13) + P(14) = 1 - P(X \leq 9) \approx$
(circle one) **0.000072 / 0.00072 / 0.0072**.
(StatCrunch: Stat, Calculators, Binomial, n: 14, p: 0.21, Prob(X $\geq$ 10) Compute.)

(e) Chance *between* 7 and 10 widgets defective, *inclusive*
$P(7 \leq X \leq 10) = P(7)+P(8)+P(9)+P(10) = P(X \leq 10) - P(X \leq 6) \approx$
(circle one) **0.005 / 0.015 / 0.034**.
(Stat, Calculators, Binomial, Between, n: 14, p: 0.21, Prob(7 $\leq$ X $\leq$ 10) Compute.)

Since $P(7 \leq X \leq 10) \approx 0.015 < 0.05$, it **is / is not**
unusual to have 7 and 10 widgets defective, inclusive.

(f) *Expected number of defectives*
$\mu_X = np = 14(0.21) =$
(circle one) **2.44 / 2.51 / 2.94**.

(g) *Standard deviation in number of defectives*:
$\sigma_X = \sqrt{np(1-p)} \approx$ (circle one) **1.52 / 1.63 / 1.76**.
So, expect 2.94 "±" 1.52 defectives.

(h) *Probability histograms, for different chance each widget defective, p.*
Identify different p for different probability histograms.

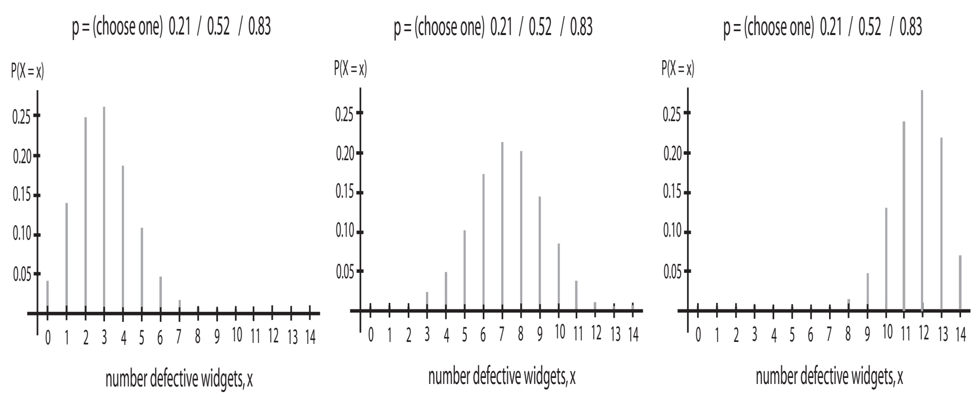

Figure 6.6 (Probability histograms: number of defective, different p.)

Notice, probability histograms for $p = 0.21$, $p = 0.52$ and $p = 0.83$ are right-skewed, (more or less) symmetric and left-skewed, respectively.

4. *Binomial: number of correct multiple choice answers.*
   On a multiple–choice exam with 4 possible choices for each of 5 questions, what is probability a student gets 3 or more correct answers just by guessing?

   (a) Since there are five questions,
   $n =$ (circle one) **3 / 4 / 5**

   (b) Student wants 3 or more correct answers,
   $x =$ (circle one) **3 / 3, 4 / 3, 4, 5**

   (c) Each question has 4 possible choices and student is choosing at random,
   $p =$ (circle one) $\mathbf{\frac{1}{3}}$ / $\mathbf{\frac{1}{4}}$ / $\mathbf{\frac{1}{5}}$

   (d) Chance student gets *at least* 3 correct answers just by guessing is
   $P(X \geq 3) = P(3) + P(4) + P(5) = 1 - P(X \leq 2) \approx$
   (choose one) **0.097 / 0.104 / 0.112**
   (StatCrunch: Stat, Calculators, Binomial, n: 5, p: 0.25, Prob(X ≥ 3) Compute.)

   Since $P(X \geq 3) \approx 0.104 > 0.05$, it **is / is not** unusual to get at least 3 correct answers just by guessing.

   (e) *Expected number of correct answers*
   $\mu_X = np = 5 \times \frac{1}{4} =$
   (circle one) $\mathbf{\frac{1}{4}}$ / $\mathbf{\frac{3}{4}}$ / $\mathbf{\frac{5}{4}}$ $= 1.25$.

   (f) *Standard deviation in number of correct answers:*
   $\sigma_X = \sqrt{np(1-p)} = \sqrt{5 \times \frac{1}{4}\left(1 - \frac{1}{4}\right)} \approx$ (circle one) **0.52 / 0.97 / 1.06**.
   So, expect 1.25 "±" 0.97 correct answers.

   (g) *Probability histogram, large n.*
   As n increases; specifically, for $np(1-p) \geq 10$, *any* binomial is symmetric, in fact, *bell*-shaped enough so empirical rule[2] can be used. Since

   $$np(1-p) = 5 \times 0.25 \times (1 - 0.25) = 0.9375 < 10$$

   probability histogram **is / is not** bell-shaped enough to use empirical rule.

6.3 The Poisson Probability Distribution

Poisson probability function, used to describe probabilities of count, X, of occurrences (successes) of an event in a specified time (or space) interval, is

$$P(x) = e^{-\lambda t}\frac{(\lambda t)^x}{x!}, \; x = 0, 1, \ldots,$$

[2]Recall, empirical rule says 68%, 95% and 99.7% inside $\mu \pm \sigma$, $\mu \pm 2\sigma$ and $\mu \pm 3\sigma$, respectively. Empirical rule allows us to say, for example, observation is *un*usual if *out*side $\mu \pm 2\sigma$.

where $\lambda > 0$ is *average* number of occurrences of event in interval of length 1, $e = 2.71828\ldots$ and expected value, variance and SD are[3],

$$\mu_X = \lambda t; \quad \sigma_X^2 = \lambda t; \quad \sigma_X = \sqrt{\lambda t} = \sqrt{\mu_X}$$

Random variable X obeys conditions of Poisson process:

- zero probability of two or more successes in sufficiently small subinterval[4],
- probability of success same if two intervals are of equal length,
- number successes in one interval independent of number of successes in any other nonoverlapping interval.

Exercise 6.3 (The Poisson Probability Distribution)

1. *Poisson: number of photon hits.*
   Piece of iron bombarded with electrons releases a number of photons, X, to a surrounding magnetic detection field. An average of $\lambda = 5$ particles hit magnetic detection field *per (one) microsecond.* Assume this is a Poisson process.

   (a) Chance $x = 0$ photons hit field in one microsecond

   $$P(0) = e^{-\lambda t}\frac{(\lambda t)^x}{x!} = \frac{e^{-(5\cdot 1)}\,(5\cdot 1)^0}{0!} \approx$$

   (circle one) **0.007** / **0.008** / **0.009**.
   (StatCrunch: Stat, Calculators, Poisson, Mean: 5, Prob(X = 0) Compute.)

   (b) Chance $x = 2$ photons hit field in one microsecond

   $$P(2) = \frac{e^{-(5\cdot 1)}\,(5\cdot 1)^2}{2!} \approx$$

   (circle one) **0.06** / **0.07** / **0.08**.
   (StatCrunch: Stat, Calculators, Poisson, Mean: 5, Prob(X = 2) Compute.)

[3]Since $\mu = \lambda t$, another version of Poisson probability function is

$$P(x) = e^{-\mu}\frac{(\mu)^x}{x!},\ x = 0, 1, \ldots,$$

[4]In other words, if we make subinterval small enough, only one occurrence (success) is possible in this subinterval. We want to be sure successes do not occur on top of one another, that they are separated from one another.

(c) Chance *at most* $x = 2$ photons hit field in one microsecond

$$P(X \leq 2) = \frac{e^{-(5\cdot 1)}\,(5\cdot 1)^0}{0!} + \frac{e^{-(5\cdot 1)}\,(5\cdot 1)^1}{1!} + \frac{e^{-(5\cdot 1)}\,(5\cdot 1)^2}{2!} \approx$$

(circle one) **0.08** / **0.12** / **0.18**.

(StatCrunch: Stat, Calculators, Poisson, Mean: 5, Prob(X $\leq$ 2) Compute.)

(d) Chance *at least* $x = 2$ photons hit field in one microsecond

$$\begin{aligned} P(X \geq 2) &= \frac{e^{-(5\cdot 1)}\,(5\cdot 1)^2}{2!} + \frac{e^{-(5\cdot 1)}\,(5\cdot 1)^3}{3!} + \cdots \text{ forever} \\ &= 1 - P(X < 2) \;=\; 1 - P(X \leq 1) \\ &= 1 - [P(0) + P(1)] \\ &= 1 - \left[\frac{e^{-(5\cdot 1)}\,(5\cdot 1)^0}{0!} + \frac{e^{-(5\cdot 1)}\,(5\cdot 1)^1}{1!}\right] \approx \end{aligned}$$

(circle one) **0.91** / **0.93** / **0.96**.

(StatCrunch: Stat, Calculators, Poisson, Mean: 5, Prob(X $\geq$ 2) Compute.)

(e) Chance $x = 2$ photons hit field in *two* microseconds

$$P(2) = e^{-\lambda t}\frac{(\lambda t)^x}{x!} = \frac{e^{-(5\cdot 2)}\,(5\cdot 2)^2}{2!} \approx$$

(circle one) **0.002** / **0.005** / **0.006**.

(Since $\mu = 5 \times 2 = 10$, Stat, Calculators, Poisson, Mean: 10, Prob(X = 2) Compute.)

(f) Chance *at most* $x = 3$ photons hit field in *two* microseconds

$$P(X \leq 3) = \frac{e^{-(5\cdot 2)}\,(5\cdot 2)^0}{0!} + \frac{e^{-(5\cdot 2)}\,(5\cdot 2)^1}{1!} + \frac{e^{-(5\cdot 2)}\,(5\cdot 2)^2}{2!} + \frac{e^{-(5\cdot 2)}\,(5\cdot 2)^3}{3!} \approx$$

(circle one) **0.01** / **0.06** / **0.08**.

(Since $\mu = 5 \times 2 = 10$, Stat, Calculators, Poisson, Mean: 10, Prob(X $\leq$ 3) Compute.)

(g) Chance *at least* $x = 21$ photons hit field in *four* microseconds

$$P(X \geq 21) = 1 - P(X < 21) = 1 - P(X \leq 20) \approx$$

(circle one) **0.44** / **0.52** / **0.68**.

(Since $\mu = 5 \times 4 = 20$, Stat, Calculators, Poisson, Mean: 20, Prob(X $\geq$ 21) Compute.)

(h) *Expected number of photon hits in one microsecond.*
$\mu = \lambda t = 5 \times 1 =$ (circle one) **3** / **4** / **5**.

(i) *Standard deviation in number of photon hits in one microsecond.*
$\sigma = \sqrt{\lambda t} = \sqrt{5 \times 1} \approx$ (circle one) **2.13** / **2.24** / **2.45**.
So, expect 5 "±" 2.24 hits in one microsecond.

(j) *Expected number of photon hits in three microseconds.*
$\mu = \lambda t = 5 \times 3 =$ (circle one) **10** / **15** / **20**.

(k) *Standard deviation in number of photon hits in three microsecond.*
$\sigma = \sqrt{\lambda t} = \sqrt{5 \times 3} \approx$ (circle one) **3.13** / **3.24** / **3.87**.
So, expect 15 "±" 3.87 hits in three microseconds.

(l) *Poisson process?* Match columns.

Poisson process	photon example
(a) zero chance more than one success in small subinterval	(A) *assume* "zero" chance two hits at same time
(b) chance of success same if two intervals are of equal length	(B) *assume* chance number hits same per microsecond
(c) independence of successes in different intervals	(C) *assume* hits independent of one another

Poisson process	(a)	(b)	(c)
photon example			

2. *Poisson: bryozoan count.*
During a biology study, 250 bryozoans are found attached on a submerged 10 centimeter by 10 centimeter (100 centimeters^2) plate, an average of $\lambda = \frac{250}{100} =$ 2.5 bryozoans per one centimeter^2. Assume this is a Poisson process.

(a) Chance $x = 0$ bryozoans attached in one centimeter^2

$$P(0) = e^{-\lambda t}\frac{(\lambda t)^x}{x!} = \frac{e^{-(2.5\cdot 1)}\,(2.5\cdot 1)^0}{0!} \approx$$

(circle one) **0.07** / **0.08** / **0.09**.
(Stat, Calculators, Poisson, Mean: 2.5, Prob(X = 0) Compute.)

(b) Chance *at most* $x = 2$ bryozoans attached in one centimeter^2

$$P(X \le 2) = \frac{e^{-(2.5\cdot 1)}\,(2.5\cdot 1)^0}{0!} + \frac{e^{-(2.5\cdot 1)}\,(2.5\cdot 1)^1}{1!} + \frac{e^{-(2.5\cdot 1)}\,(2.5\cdot 1)^2}{2!} \approx$$

(circle one) **0.54** / **0.62** / **0.78**.
(Stat, Calculators, Poisson, Mean: 2.5, Prob(X ≤ 2) Compute.)

(c) Chance *between* $x = 2$ and $x = 4$ bryozoans in one centimeter^2, inclusive:

$$P(2 \le X \le 4) = P(X \le 4) - P(X < 2) = P(X \le 4) - P(X \le 1) \approx$$

(circle one) **0.60** / **0.73** / **0.76**.
(Stat, Calculators, Poisson, Between, Mean: 2.5, Prob(2 ≤ X ≤ 4) Compute.)

(d) Chance *between* $x = 2$ and $x = 4$ bryozoans in *three* centimeters^2, inclusive:

$$P(2 \le X \le 4) = P(X \le 4) - P(X < 2) = P(X \le 4) - P(X \le 1) \approx$$

(circle one) **0.13** / **0.23** / **0.36**.
(Stat, Calculators, Poisson, Between, Mean: 7.5, Prob(2 ≤ X ≤ 4) Compute.)

(e) *Expected number of bryozoans in three centimeters*^{2}.
$\mu = \lambda t = 2.5 \times 3 =$ (circle one) **7.5** / **10** / **12.5**.

(f) *Standard deviation in number of bryozoans in three centimeters*^{2}.
$\sigma = \sqrt{\lambda t} = \sqrt{2.5 \times 3} \approx$ (circle one) **2.74** / **3.24** / **3.87**.
So, expect 7.5 "±" 2.74 bryozoans in three centimeters^2.