

**Discrete Probability Distributions**

[Navidi sections 4.1-4.4; Devore sections 3.4-3.6]

Chapter 5 introduced the concept of probability mass functions for discrete random quantities. The only standard distribution introduced then is the **discrete uniform distribution**, whose p.m.f. is

$$p(x) = \frac{1}{n} \quad (x = x_1, x_2, \dots, x_n)$$

This chapter will introduce some more standard probability mass functions.

---

**Bernoulli Probability Distribution**

Where the outcome of a single trial can take one of only two possible values, “success” or “failure”, the Bernoulli random quantity maps “success” to  $X = 1$  and “failure” to  $X = 0$ . “Success” and “failure” are a complementary pair of simple events. “Success” can mean “heads” on one flip of a coin, or “defective” when a single item is tested, or “total is 7” on one roll of two dice, etc. If the probability of a success is  $p$ , then the p.m.f. follows immediately:

$$p(x) = \begin{cases} p & (x=1) \\ 1-p & (x=0) \\ 0 & (\text{otherwise}) \end{cases}$$

The mean and variance are calculated easily:

$$E[X] =$$

and

$$E[X^2] =$$

$$\Rightarrow V[X] = E[X^2] - (E[X])^2 =$$

---

Example 9.01:

Let  $X = 1$  correspond to “total score on one roll of a pair of fair standard six-sided dice is 7” and  $X = 0$  in the event of some other total, then  $X$  has a Bernoulli distribution. Find the parameter  $p$  of this distribution and find the mean and variance of  $X$ .

---

**The Binomial Distribution**

The most important discrete probability distribution is the binomial distribution. Here we develop the probability mass function through an example.

Example 9.02

It is known that each component, as it emerges from an experimental production process, has a probability of 20% to be defective (and therefore a probability of 80% to be good). A random sample of ten components is drawn.

- (a) Find the probability that exactly three components in the random sample are defective.

---

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

$S$  = success (the component is defective) and

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

[It may seem strange to associate “success” with “defective”, but it is the number of defective items in which we are interested.]

This is a Bernoulli trial (a pair of complementary outcomes only), with constant  $P[S]$ .

Also let

$E$  = the desired event (three defective and seven good components in the random sample)

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

---

Example 9.02(a) (continued)

However, inside the event  $E$ , the three successes may occur anywhere among the ten trials. The number of distinct rearrangements of three successes in ten trials is

Therefore  $P[E] =$

---

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

$$P[S^x F^{10-x}] =$$

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

Therefore the probability mass function of  $X$  is

$$P[\text{exactly } x \text{ successes}] = P[X = x] = \quad (x = 0, 1, 2, \dots, 10)$$

---

We can generalize this result further. Let the proportion of defective items in the population be  $p$  (rather than 20%) and let the sample size be  $n$  (rather than 10). Then the **binomial probability mass function** becomes

$$P[X = x] = b(x; n, p) =$$

In this example, we have drawn a random sample from a conceptual population, so that every trial is independent of every other trial.

---

**Conditions for a discrete random quantity to have a binomial p.m.f.:**

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 all hold:

- (1) Each trial has exactly two complementary outcomes (“success” and “failure”);
- (2) The [unconditional] 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.

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 9.03

Show that condition (3) is not satisfied if a random sample of size 2 is taken from a population of size 5 with 40% successes. Show that condition (3) is almost satisfied if a random sample of size 2 is taken from a population of size 5000 with 40% successes.

$S \quad S \quad F \quad F \quad F$

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

then  $P[S_2] = 2/5 = .4000$

but  $P[S_2 | S_1] =$

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  $P[S_2] = 2,000 / 5,000 = .4000$

but  $P[S_2 | S_1] =$

Therefore the outcomes of the two trials are nearly independent and the binomial p.m.f. may be used as an excellent approximation to the true probability distribution.

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 Navidi (and Devore), table A.1, for  $n$  up to 20 and values of  $p$  that are integer multiples of .05. It can be evaluated for any valid choice of  $(n, p)$  using the file at "[www.engr.mun.ca/~ggeorge/4421/demos/Binomial.xls](http://www.engr.mun.ca/~ggeorge/4421/demos/Binomial.xls)".

---

#### Example 9.04

The probability mass function of the random quantity  $X$  is known to be binomial with parameters  $n = 10$  and  $p = .20$ . Find  $P[1 < X < 5]$ .

---

Example 9.04 (continued)

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/4421/demos/Binomial.xls](http://www.engr.mun.ca/~ggeorge/4421/demos/Binomial.xls)".

To find the values of  $B(4; 10, .20)$  and  $B(1; 10, .20)$ , the following partial table of values for  $b(x; 10, .20)$  and  $B(x; 10, .20)$ , (drawn from the Excel file), may be used:

$n = 10$		$p = 0.2$	
<b>p.m.f.</b>		<b>c.d.f.</b>	
$x$	$P[X = x]$	$x$	$P[X \leq x]$
0	0.10737	0	0.10737
1	0.26844	1	0.37581
2	0.30199	2	0.67780
3	0.20133	3	0.87913
4	0.08808	4	0.96721
5	0.02642	5	0.99363
6	0.00551	6	0.99914
7	0.00079	7	0.99992
8	0.00007	8	1.00000
9	0.00000	9	1.00000
10	0.00000	10	1.00000

If a question involving the calculation of several adjacent binomial probability masses arises in a test or examination, then the appropriate table of values, similar to that shown above, will be provided with the question paper.

---

Example 9.05

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 (at least to an excellent approximation).  $P[X = x] = b(x; 20, .10)$ .

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

Example 9.05 (continued)

To find the value of  $B(2; 20, .10)$ , the partial table of values for  $b(x; 20, .10)$  and  $B(x; 20, .10)$  may be used:

$n = 20$		$p = 0.1$	
<b>p.m.f.</b>		<b>c.d.f.</b>	
$x$	$P[X = x]$	$x$	$P[X \leq x]$
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  $X$ ,

$$E[X] = np \quad \text{and} \quad V[X] = np(1-p)$$

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

---



---

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  $E[X]$ .

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

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

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

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

Example 9.06

Find the probability of

- (a) exactly one '5' or '6' when four dice are thrown.
  - (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  $P[X = x] = b(x; 4, 1/3)$ .

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

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

(c)  $E[X] =$

---

---

Example 9.07

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  $P[X = x] = b(x; n, .02)$ .

---

$$P[X > 0] =$$

As a check,

$$\begin{aligned} n = 148 &\Rightarrow P[X > 0] = 1 - (.98)^{148} \approx 94.97\% \\ n = \mathbf{149} &\Rightarrow P[X > 0] = 1 - (.98)^{149} \approx \mathbf{95.07\%} . \end{aligned}$$

---

### Estimation of Population Proportion

In many cases we do not know the value of the success probability  $p$  (which is also the proportion of successes in the entire population). A point estimate of an unknown value of  $p$  may be found by conducting  $n$  independent Bernoulli trials – that is, take a random sample of size  $n$  from the population and for each item record whether it is a success or a failure. The sample proportion  $\hat{P}$  is a random quantity that is an **estimator** for  $p$ :

$$\hat{P} = \frac{\text{number of successes}}{\text{number of trials}} = \frac{X}{n}$$

An actual observed sample proportion is an **estimate** for  $p$ :  $\hat{p} = \frac{x}{n}$ .

The random quantity  $X$  is binomial with parameters  $(n, p)$

$$\Rightarrow E[\hat{P}] =$$

Therefore the sample proportion  $\hat{P}$  is an unbiased estimator of the population proportion  $p$ .

$$V[\hat{P}] =$$

The uncertainty  $\sigma$  in  $\hat{P}$  is

Since  $p$  is the very parameter that we are trying to estimate, it and  $\sigma$  are unknown.

The simplest solution is to replace the unknown  $p$  by its point estimate  $\hat{p} = \frac{x}{n}$  in the

expression for the uncertainty:

[Later we shall see a better estimate.]

#### Example 9.08

Suppose that there are 13 successes in a random sample of size 20 drawn from a much larger population. Then our estimate of the true proportion of successes in the population is

$$p \approx$$

## Hypergeometric Probability Distribution

This distribution will also be developed by way of an example.

### Example 9.09

A truck carries nine crates of printer paper and three crates of ring binders. Four crates are taken at random from the truck. Find the probability that at least two of the four crates contain ring binders.

---

Let  $X$  = the number of crates of ring binders in the random sample of four crates.  
 $P[X \geq 2]$  is required.

Let us check the four conditions for a binomial distribution:

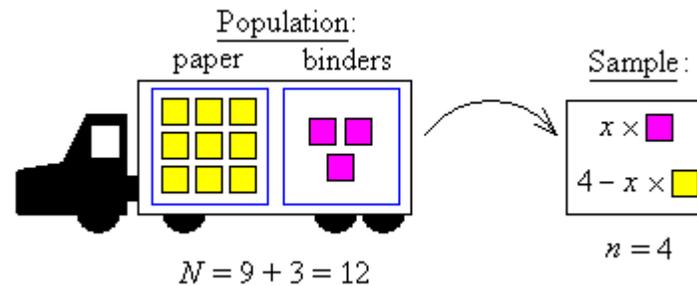
- Complementary pair of outcomes in each trial:  
Each crate either contains ring binders (= “success”) or printer paper ✓  
(=“failure”)
- $P[S] = \text{constant}$   
For each crate in the sample,  $P[S] = 3/12$  ✓  
(when you don’t know the contents of the other crates).
- Trials are independent
  
- Sample size is constant ✓  
 $n = 4$

Therefore the binomial p.m.f. cannot be used (not even as an approximation).

Examine the population and the sample (next page):

---

Example 9.09 (continued)



The number of equally likely sample points in the sample space is the number of ways of drawing the random sample of four crates from the twelve crates on the truck:

$$n(S) =$$

The number of ways in which exactly  $x$  crates of binders can appear among the four crates in the random sample is

$$n(E) =$$

The probability mass function for  $X$  follows:

The general case for the hypergeometric p.m.f. follows.

If a population of finite size  $N$  contains  $R$  successes and  $(N - R)$  failures, a random sample of size  $n$  is drawn from this population, and  $X$  represents the number of successes in the sample, then

$$P[X = x] = p(x) =$$

This probability mass is zero unless all of the following are true:

$n$  is an integer in the range  $[0, N]$  and  
 $x$  is an integer in the range  $[0, R]$  and  
 $(n - x)$  is an integer in the range  $[0, N - R]$ .

This distribution can be obtained from first principles, using counting techniques, as in Example 9.09.

An Excel spreadsheet for this distribution is available at  
["www.engr.mun.ca/~ggeorge/4421/demos/Hypergeom.xls"](http://www.engr.mun.ca/~ggeorge/4421/demos/Hypergeom.xls).

*Non-examinable*; for reference only:

The mean and variance of a random quantity that follows the hypergeometric distribution,  $X \sim H(n, R, N)$  are

$$\mu = E[X] = n \frac{R}{N} \quad \text{and} \quad \sigma^2 = V[X] = n \frac{R}{N} \left(1 - \frac{R}{N}\right) \left(\frac{N - n}{N - 1}\right)$$

In the limit as  $R$  and  $N$  go to infinity, such that  $p = \frac{R}{N}$  remains constant, the hypergeometric distribution becomes the binomial distribution  $b(n, p)$ , with  $\mu = np$  and  $\sigma^2 \rightarrow np(1 - p)$ .

As a general rule of thumb, the binomial distribution may be used as an acceptably good approximation to the hypergeometric distribution if the sample size  $n$  is no more than 5% of the population size  $N$ .

**Negative Binomial Probability Distribution**

This distribution will also be developed by way of an example.

Example 9.10:

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  $E =$

so  $P[E] =$

---

This is an example of the negative binomial distribution  $NB(x; r, p)$ ,  
(Navidi, section 4.4, pages 234-237; Devore, section 3.5, pages 118-120).  
In this case, with  $X =$  (number of attempts needed to obtain the second success),  
 $X \sim NB(x; 2, .8)$  and  $P[E] = P[X \leq 3] = P[X = 2] + P[X = 3]$ .

---

The p.m.f., mean and variance of the negative binomial random quantity  $X \sim \text{NB}(x; r, p)$  are

$$P[X = x] = p(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, \quad (x = r, r+1, r+2, \dots)$$

$$\mu = E[X] = \frac{r}{p}, \quad \sigma^2 = V[X] = \frac{r(1-p)}{p^2}$$

A special case of the negative binomial distribution occurs when  $r = 1$ : the **geometric distribution**, where

$X =$  the number of trials up to and including the *first* success  $\sim \text{Geom}(p)$ .

$$P[X = x] = p(x) = p(1-p)^{x-1}, \quad (x \in \mathbb{N})$$

$$\mu = E[X] = \frac{1}{p}, \quad \sigma^2 = V[X] = \frac{1-p}{p^2}$$

Knowledge of the formulae on this page will *not* be assumed in a test or the final examination. They are derived in the Navidi and Devore textbooks. Some practice is provided in the problem sets.

---

---

### Poisson Distribution

If the discrete random quantity  $X$  is a count of the number of times an event occurs in some fixed interval of time (or length or area or volume), then the probability distribution for  $X$  may be Poisson. This situation can arise if the waiting time  $T$  from one occurrence to the next is independent of all previous waiting times (see the exponential distribution in Chapter 10).

The Poisson distribution can also arise as a special case of the binomial distribution, in the limit as the sample size increases to infinity while the expected number of successes,  $\mu = np$ , is held

constant. Let  $\lambda = np \Rightarrow p = \frac{\lambda}{n}$ , then the binomial p.m.f. becomes

$$\begin{aligned} P[X = x] &= \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} = \frac{n(n-1)\dots(n-(x-1))}{x!} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} \\ &= \frac{n}{n} \times \frac{n-1}{n} \times \dots \times \frac{n-(x-1)}{n} \frac{\lambda^x}{x!} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-x} \\ &= 1 \times \left(1 - \frac{1}{n}\right) \times \dots \times \left(1 - \frac{x-1}{n}\right) \frac{\lambda^x}{x!} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-x} \\ &\Rightarrow \lim_{\substack{n \rightarrow \infty \\ \lambda = \text{const.}}} P[X = x] = 1 \times 1 \times \dots \times 1 \times \frac{\lambda^x}{x!} \times \lim_{\substack{n \rightarrow \infty \\ \lambda = \text{const.}}} \left(1 - \frac{\lambda}{n}\right)^n \times (1-0)^{-x} \end{aligned}$$

From ENGI 3425 (or equivalent), we know that  $\lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n = e^x$

Therefore

$$\lim_{\substack{n \rightarrow \infty \\ \lambda = \text{const.}}} b\left(x; n, \frac{\lambda}{n}\right) = \lim_{\substack{n \rightarrow \infty \\ \lambda = \text{const.}}} {}^n C_x p^x (1-p)^{n-x} = \frac{\lambda^x e^{-\lambda}}{x!}$$

The parameter  $\lambda$  is the mean number of successes in the fixed interval.

The Poisson probability mass function is

$$P[X = x] = \text{Poisson}(x; \lambda) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad (x = 0, 1, 2, 3, \dots)$$

We shall see later that, for a Poisson random quantity  $X$ ,  $\sigma^2 = \mu = \lambda$ .

Example 9.11

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

- (a) there is at least one crack in a one metre long cable.
  - (b) there are less than 20 cracks in a five metre long cable.
- 

(a) Let  $X =$  (number of cracks in the one-metre long cable)

$X \sim \text{Poisson}(3.4)$

$$P[X \geq 1] =$$

(b) Let  $Y =$  (number of cracks in the five-metre long cable)

If, on average, there are 3.4 cracks in every one metre, then there are  $5 \times 3.4 = 17$  cracks on average in every five metres.

$Y \sim \text{Poisson}(17)$

$$P[Y < 20] =$$

---

Example 9.11 (continued)

The following extracts from the Excel spreadsheet file

"www.engr.mun.ca/~ggeorge/4421/demos/Poisson.xls" can be used instead.

The value of the parameter  $\lambda$  of the Poisson distribution is  $\lambda = 17$  in part (b).

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

$\mu = 17$			
<b>p.m.f.</b>		<b>c.d.f.</b>	
$x$	$P[X = x]$	$x$	$P[X \leq x]$
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

If a question involving the calculation of several adjacent Poisson probability masses arises in a test or examination, then the appropriate table of values, similar to that shown above, will be provided with the question paper.

---

### Mean and Variance of a Poisson Random Quantity

This derivation is *not examinable*, but the result is quotable.

If  $P[X = x] = \frac{\lambda^x}{x!} e^{-\lambda}$ , then  $\sum_{x=0}^{\infty} \frac{\lambda^x}{x!} e^{-\lambda} = 1$  (coherence condition) and

$$E[X] = \sum_{x=0}^{\infty} x \cdot \frac{\lambda^x}{x!} e^{-\lambda} = 0 + \sum_{x=1}^{\infty} x \cdot \frac{\lambda^x}{x!} e^{-\lambda} = \sum_{x=1}^{\infty} \frac{\lambda \cdot \lambda^{x-1}}{(x-1)!} e^{-\lambda}$$

Let  $y = x - 1$ , then  $E[X] = \lambda \sum_{y=0}^{\infty} \frac{\lambda^y}{y!} e^{-\lambda} = \lambda \times 1 = \lambda$

---

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

Let  $t = x - 2$ , then

$$E[X^2] = \lambda^2 \sum_{t=0}^{\infty} \frac{\lambda^t}{t!} e^{-\lambda} + \lambda = \lambda^2 + \lambda$$

$$V[X] = E[X^2] - (E[X])^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

Therefore

$$\sigma^2 = \mu = \lambda$$


---

Example 9.12 (Navidi textbook, exercises 4.3, page 228, question 10)

A chemist wishes to estimate the concentration of particles in a certain suspension. She withdraws 3 mL of the suspension and counts 48 particles. Estimate the concentration in particles per mL and find the uncertainty in the estimate.

---

We assume that the concentration  $c$  (number of particles per mL) is constant across all volumes. Then the number  $N$  of particles in volume  $v$  has the distribution  $N \sim \text{Poisson}(\lambda)$ , where  $\lambda = cv$ .

We do not know the true concentration  $c$  with certainty.

We do have a point estimate,

---

Example 9.13

Ships arrive at a port on average every 30 minutes, independently of all previous arrivals.

No two ships ever arrive at exactly the same time.

Find the probability that, during the next hour,

- (a) no ships arrive;
  - (b) more than four ships arrive.
- 

Ships arrive, on average, at a rate  $\frac{1 \text{ ship}}{30 \text{ min}} \times \left(60 \frac{\text{min}}{\text{hr}}\right) = 2$  ships per hour.

Therefore the number of ships  $N$  that will arrive in the next hour is a random quantity that follows the distribution  $N \sim \text{Poisson}(2)$ .

(a)  $P[N = 0] =$

(b)  $P[N > 4] =$

[Space for Additional Notes]

---