Probability Theory

Textbook: A First Course in Probability, Sheldon Ross, 2019.

Prof. Hicham Elmongui

elmongui@alexu.edu.eg

Chapter 04: Discrete Random Variables

Random Variables

We are often interested in a *function* of the outcome as opposed to the actual outcome (e.g., total sum of dice faces, number of coin tosses)

Definition

Random variables are real-valued functions defined on the sample space.

Suppose that we are tossing 3 fair coins. If we let Y denote the number of heads that appear, then Y is a random variable taking on one of the values 0, 1, 2, and 3 with respective probabilities

$$P\{Y=0\} = P\{(T,T,T)\} = \frac{1}{8}$$
$$P\{Y=1\} = P\{(T,T,H),(T,H,T),(H,T,T)\} = \frac{3}{8}$$
$$P\{Y=2\} = P\{(H,H,T),(H,T,H),(T,H,H)\} = \frac{3}{8}$$
$$P\{Y=3\} = P\{(H,H,H)\} = \frac{1}{8}$$

Example

Example

Three balls are to be randomly selected without replacement from an urn containing 20 balls numbered 1 through 20. If we bet that at least one of the balls that are drawn has a number as large as or larger than 17, what is the probability that we win the bet?

- Let X denote the largest number selected
- Then X is a random variable taking on one of the values $3, 4, \cdots, 20$

$$P\{X=i\} = \frac{\binom{i-1}{2}}{\binom{20}{3}}, 3 \le i \le 20$$
$$P\{X=20\} \approx .150$$
$$P\{X=19\} \approx .134$$

$$P\{X=18\} \approx .119$$
$$P\{X=17\} \approx .105$$
$$P\{winning\} = P\{X \ge 17\}$$
$$\approx .150 + .134 + .119$$
$$+ .105 = .508$$

Discrete Random Variables

Definition

A random variable that can take on at most a countable number of possible values is said to be *discrete*.

Probability Mass Function

The *probability mass function* of a discrete random variable $X \triangleq$

$$p(a) = P\{X = a\}$$

If X must assume one of the values $x_1, x_2, \cdots$, then

$$p(x_i) \ge 0 \quad \text{for } i = 1, 2, \cdots$$
$$p(x) = 0 \quad \text{for all other values of } x$$
$$\sum_{i=1}^{\infty} p(a) = 1$$

Example

Example

The pmf of a random variable X is given by $p(i) = c\lambda^i/i!, i = 0, 1, 2, \cdots,$ where λ is some positive value. Find

1. $P\{X = 0\}$
2. $P\{X > 2\}$

$$1 = \sum_{i=0}^{\infty} p(i) = c\sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = ce^{\lambda} \qquad \Rightarrow c = e^{-\lambda}$$

1. $P\{X=0\} = e^{-\lambda}\frac{\lambda^0}{0!} = e^{-\lambda}$
2. $P\{X>2\} = 1 - P\{X \le 2\}$
   $= 1 - P\{X=0\} - P\{X=1\} - P\{X=2\}$
   $= 1 - e^{-\lambda} - \lambda e^{-\lambda} - \frac{\lambda^2 e^{-\lambda}}{2}$

Graphical Representation of a PMF

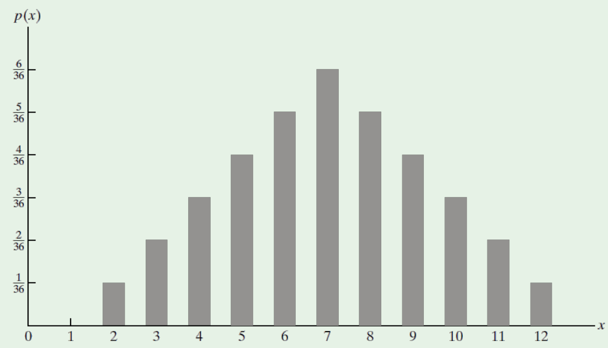


Figure: A graph of the probability mass function of the random variable representing the sum when two dice are rolled.

Cumulative Distribution Function

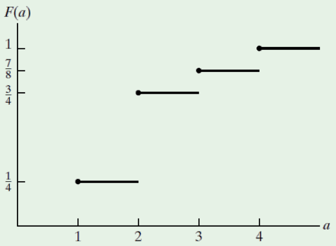
The cumulative distribution function F (or simply – the distribution function) can be expressed in terms of $p(a)$ by

$$F(a) = \sum_{\text{all } x \leq a} p(x)$$

If X has a probability mass function given by

$p(1) = \frac{1}{4}$ $p(2) = \frac{1}{2}$ $p(3) = \frac{1}{8}$ $p(4) = \frac{1}{8}$

$$F(a) = \begin{cases} 0 & a < 1 \\ \frac{1}{4} & 1 \leq a < 2 \\ \frac{3}{4} & 2 \leq a < 3 \\ \frac{7}{8} & 3 \leq a < 4 \\ 1 & 4 \leq a \end{cases}$$



Expected Value

The *expectation* (or *expected value*) of a discrete RV X

$E[X]$ is a weighted average of the possible values that X can take on, each value being weighted by the probability that X assumes it.

$$E[X] = \sum_{x:p(x)>0} xp(x)$$

Example

Find $E[X]$, where X is the outcome when we roll a fair die.

$$E[X] = 1\left(\frac{1}{6}\right) + 2\left(\frac{1}{6}\right) + 3\left(\frac{1}{6}\right) + 4\left(\frac{1}{6}\right) + 5\left(\frac{1}{6}\right) + 6\left(\frac{1}{6}\right) = \frac{7}{2}$$

Analogy with "center of gravity" of a distribution of mass

Expectation of a Function of a Random Variable

- We are given a discrete RV, X, along with its pmf
- A function $g(X)$ is itself a discrete RV
- Its pmf can be determined from the pmf of X
- we can compute $E[g(X)]$ by using the definition of expected value

Example

Compute $E[X^2]$, where X denotes a random variable that takes on any of the values -1, 0, and 1 with respective probabilities

$P\{X = -1\} = .1$ $P\{X = 0\} = .3$ $P\{X = 1\} = .6$

Let $Y = X^2$. The pmf of Y is given by

$$P\{Y = 1\} = P\{X = -1\} + P\{X = 1\} = .7$$
$$P\{Y = 0\} = P\{X = 0\} = .3$$

Hence, $E[X^2] = E[Y] = 1(.7) + 0(.3) = .7$

Expectation of a Function of a RV (cont'd)

Proposition

If X is a discrete RV that takes on one of the values $x_i, i \geq 1$, with respective probabilities $p(x_i)$, then, for any real-valued function g,

$$E[g(X)] = \sum_i g(x_i)p(x_i)$$

Proof: by grouping all terms having the same value of $g(x_i)$

$$\begin{aligned}\sum_i g(x_i)p(x_i) &= \sum_j \sum_{i:g(x_i)=y_j} g(x_i)p(x_i) \\ &= \sum_j \sum_{i:g(x_i)=y_j} y_j p(x_i) \\ &= \sum_j y_j \sum_{i:g(x_i)=y_j} p(x_i) \\ &= \sum_j y_j P(g(X) = y_j) \qquad = E[g(X)]\end{aligned}$$

Expectation of a Linear Function of a RV

If a and b are constants, then

$$E[aX + b] = aE[X] + b$$

Proof

$$\begin{aligned}E[aX + b] &= \sum_{x:p(x)>0} (ax + b)p(x) \\ &= \sum_{x:p(x)>0} axp(x) + \sum_{x:p(x)>0} bp(x) \\ &= a\sum_{x:p(x)>0} xp(x) + b\sum_{x:p(x)>0} p(x) \\ &= aE[X] + b\end{aligned}$$

The expected value of a random variable X, $E[X]$, is also referred to as the *mean* or the *first moment of* X. The quantity $E[X^n], n \geq 1$, is called the n^{th} *moment of* X.

Variance

Is the expectation enough to summarize a distribution function?

$$Y = \begin{cases} -1 & \text{with probability } \frac{1}{2} \\ +1 & \text{with probability } \frac{1}{2} \end{cases} \qquad Z = \begin{cases} -30 & \text{with probability } \frac{1}{4} \\ +10 & \text{with probability } \frac{3}{4} \end{cases}$$

The expected value of a RV tells nothing about the variation, or spread, of its possible values.

- We expect X to take on values around its mean $E[X]$
- To measuring the possible variation of X, we may look at how far apart X would be from its mean, on the average

$$E[|X - \mu|], \quad \text{where } \mu = E[X]$$

- It turns out to be mathematically inconvenient to deal with this quantity

Variance

If X is a RV with mean μ, then the variance of X, denoted by $\mathrm{Var}(X)$, is defined by

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

An alternative formula for Var(X)

$$\begin{aligned}\mathrm{Var}(X) &= E\left[(X-\mu)^2\right] \\ &= E[X^2 - 2\mu X + \mu^2] \\ &= E[X^2] - 2\mu E[X] + \mu^2 \\ &= E[X^2] - 2\mu^2 + \mu^2 = E[X^2] - \mu^2 \\ &= E[X^2] - \left(E[X]\right)^2\end{aligned}$$

Analogy: The variance represents the *moment of inertia* of a distribution of mass around its center of gravity.

Examples

Example 1

Calculate $\mathrm{Var}(X)$ if X represents the outcome when a fair die is rolled.

$$E[X^2] = 1^2\left(\tfrac{1}{6}\right) + 2^2\left(\tfrac{1}{6}\right) + 3^2\left(\tfrac{1}{6}\right) + 4^2\left(\tfrac{1}{6}\right) + 5^2\left(\tfrac{1}{6}\right) + 6^2\left(\tfrac{1}{6}\right) = \frac{91}{6}$$

$$\mathrm{Var}(X) = E[X^2] - \left(E[X]\right)^2 = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}$$

Example 2

Calculate $\mathrm{Var}(Y)$ and $\mathrm{Var}(Z)$ if

$$Y = \begin{cases} -1 & \text{with probability } \frac{1}{2} \\ +1 & \text{with probability } \frac{1}{2} \end{cases} \qquad Z = \begin{cases} -30 & \text{with probability } \frac{1}{4} \\ +10 & \text{with probability } \frac{3}{4} \end{cases}$$

$$E[Y^2] = (-1)^2\left(\tfrac{1}{2}\right) + 1^2\left(\tfrac{1}{2}\right) = 1 \qquad E[Z^2] = (-30)^2\left(\tfrac{1}{4}\right) + 10^2\left(\tfrac{3}{4}\right) = 300$$

$$\mathrm{Var}(Y) = 1 - 0^2 = 1 \qquad \mathrm{Var}(Z) = 300 - 0^2 = 300$$

Variance of a Linear Function of a RV

If a and b are constants, then

$$\mathrm{Var}(aX+b) = a^2\mathrm{Var}(X)$$

Proof

Let $\mu = E[X]$ and $Y = aX + b$, then

$$E[Y] = E[aX+b] = aE[X] + b = a\mu + b$$

Therefore,

$$\begin{aligned}\mathrm{Var}(aX+b) &= \mathrm{Var}(Y) \\ &= E\left[(Y - E[Y])^2\right] \\ &= E\left[((aX+b) - (a\mu+b))^2\right] \\ &= E\left[(aX - a\mu)^2\right] = E\left[a^2(X-\mu)^2\right] \\ &= a^2E\left[(X-\mu)^2\right] = a^2\mathrm{Var}(X)\end{aligned}$$

Standard Deviation of a RV

Definition

The square root of the $\mathrm{Var}(X)$ is called the *standard deviation* of X, and we denote it by $\mathrm{SD}(X)$.

$$\mathrm{SD}(X) = \sqrt{\mathrm{Var}(X)}$$

Expected Value of Sums of RVs

Sum of RVs

- Consider a probability experiment whose sample space S is either finite or countably infinite
- For a RV X, let $X(s)$ denote the value of X when $s \in S$ is the outcome of the experiment
- If X and Y are both RVs, then so is their sum — $Z = X + Y$

$$Z(s) = X(s) + Y(s)$$

- Suppose that the experiment consists of flipping a coin 5 times
- Suppose X is the number of heads in the first 3 flips
- Suppose Y is the number of heads in the final 2 flips
- Let $Z = X + Y$. For the outcome $s = (h, t, h, t, h)$,

$$X(s) = 2 \qquad Y(s) = 1 \qquad Z(s) = X(s) + Y(s) = 3$$

Expected Value of Sums of RVs (cont'd)

Proposition

$E[X]$ equals a weighted average of the values $X(s)$, $s \in S$, with $X(s)$ weighted by the probability that s is the outcome of the experiment

$$\sum_{s\in S} X(s)p(s)$$

Proof

Suppose that the distinct values of X are $x_i, i \geq 1$. For each i, let S_i be the event that X is equal to x_i. That is, $S_i = \{s : X(s) = x_i\}$. Then,

$$\begin{aligned}E[X] &= \sum_i x_i P\{X = x_i\} \\ &= \sum_i x_i P(S_i) \qquad &= \sum_i x_i \sum_{s\in S_i} p(s) \\ &= \sum_i \sum_{s\in S_i} x_i p(s) \qquad &= \sum_{s\in S} X(s)p(s)\end{aligned}$$

where the final equality follows because $S_1, S_2, \cdots$ are mutually exclusive events whose union is S.

Expected Value of Sums of RVs (cont'd)

Corollary

For random variables $X_1, X_2, \cdots, X_n$,

$$E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]$$

Proof

Let $Z = \sum_{i=1}^{n} X_i$. Then,

$$\begin{aligned} E[Z] &= \sum_{s\in S} Z(s)p(s) \\ &= \sum_{s\in S} \Big(X_1(s) + X_2(s) + \cdots + X_n(s)\Big)p(s) \\ &= \sum_{s\in S} X_1(s)p(s) + \sum_{s\in S} X_2(s)p(s) + \cdots + \sum_{s\in S} X_n(s)p(s) \\ &= E[X_1] + E[X_2] + \cdots + E[X_n] \end{aligned}$$

Properties of the Cumulative Distribution Function

For the distribution function F of X,

- F is a nondecreasing function; that is, if $a < b$, then $F(a) \le F(b)$
- $\lim_{b\to\infty} F(b) = 1$
- $\lim_{b\to-\infty} F(b) = 0$
- F is right continuous. That is, for any b and any decreasing sequence b_n, $n \ge 1$, that converges to b, $\lim_{n\to\infty} F(b_n) = F(b)$

Answering Probability Questions in Terms of F

$P\{a < X \le b\}$ $\quad \forall a < b$

$$\begin{aligned} \{X \le b\} &= \{X \le a\} \cup \{a < X \le b\} \\ P\{X \le b\} &= P\{X \le a\} + P\{a < X \le b\} \\ F(b) &= F(a) + P\{a < X \le b\} \\ P\{a < X \le b\} &= F(b) - F(a) \end{aligned}$$

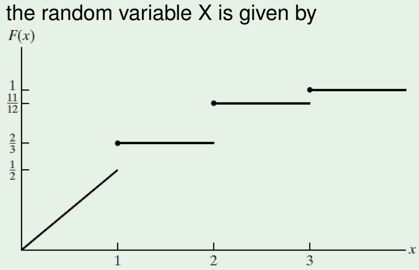
$P\{X < b\}$

$$\begin{aligned} P\{X < b\} &= P\left(\lim_{n\to\infty}\left\{X \le b - \frac{1}{n}\right\}\right) \\ &= \lim_{n\to\infty} P\left(X \le b - \frac{1}{n}\right) \\ &= \lim_{n\to\infty} F\left(b - \frac{1}{n}\right) \end{aligned}$$

Example

The distribution function of the random variable X is given by

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{x}{2} & 0 \le x < 1 \\ \frac{2}{3} & 1 \le x < 2 \\ \frac{11}{12} & 2 \le x < 3 \\ 1 & 3 \le x \end{cases}$$



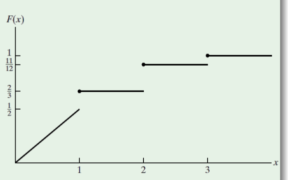
Compute

1. $P\{X < 3\}$
2. $P\{X = 1\}$
3. $P\{X > \frac{1}{2}\}$
4. $P\{2 < X \le 4\}$

Example (cont'd)

Solution

1. $P\{X < 3\} = \lim_{n\to\infty} F\left(3 - \frac{1}{n}\right) = \frac{11}{12}$
2. $P\{X = 1\} = P\{X \le 1\} - P\{X < 1\}$
   $= F(1) - \lim_{n\to\infty} F\left(1 - \frac{1}{n}\right)$
   $= \frac{2}{3} - \frac{1}{2} = \frac{1}{6}$
3. $P\{X > \frac{1}{2}\} = 1 - P\{X \le \frac{1}{2}\}$
   $= 1 - F(\frac{1}{2})$
   $= 1 - \frac{1}{4} = \frac{3}{4}$
4. $P\{2 < X \le 4\} = F(4) - F(2)$
   $= 1 - \frac{11}{12} = \frac{1}{12}$



Common Discrete Probability Distributions

- Bernoulli Distribution
- Binomial Distribution
- Poisson Distribution
- Geometric Distribution

Bernoulli Distribution

The Bernoulli Random Variable

- Suppose that a trial, or an experiment, whose outcome can be classified as either a *success* or a *failure* is performed
- A *Bernoulli RV*, X, with parameter p has the following pmf

$$p(1) = P\{X = 1\} = p$$
$$p(0) = P\{X = 0\} = q = 1 - p$$

where $p, 0 \le p \le 1$, is the probability that the trial is a success

$$\begin{aligned} E[X] &= 0 \times (1-p) + 1 \times p \\ &= p \\ E[X^2] &= 0^2 \times (1-p) + 1^2 \times p \\ &= p \end{aligned}$$

$$\begin{aligned} \mathrm{Var}(X) &= E[X^2] - \left(E[X]\right)^2 \\ &= p - p^2 \\ &= p(1-p) \\ &= pq \end{aligned}$$

Binomial Distribution

The Binomial Random Variable

- Suppose that n independent trials, each of which is a Bernoulli trial with parameter p, are to be performed.
- If X represents the number of successes that occur in the n trials, then X is said to be a *binomial RV* with parameters (n, p). Its pmf is

$$p(i) = \binom{n}{i} p^i (1-p)^{n-i} \qquad i = 0, 1, \cdots, n$$

$$\begin{aligned} \sum_{i=0}^{\infty} p(i) &= \sum_{i=0}^{n} \binom{n}{i} p^i (1-p)^{n-i} \\ &= \left(p + (1-p)\right)^n \\ &= 1^n \\ &= 1 \end{aligned}$$

Binomial Distribution

It is known that screws produced by a certain company will be defective with probability .01, independently of each other. The company sells the screws in packages of 10 and offers a money-back guarantee that at most 1 of the 10 screws is defective. What proportion of packages sold must the company replace?

Solution

If X is the number of defective screws in a package, then X is a binomial RV with parameters (10, .01).

$$\begin{aligned} P\{X > 1\} &= 1 - P\{X = 0\} - P\{X = 1\} \\ &= 1 - \binom{10}{0}(.01)^0(.99)^{10} - \binom{10}{1}(.01)^1(.99)^9 \\ &= .004 \end{aligned}$$

Thus, only .4 percent of the packages will have to be replaced.

Binomial Distribution

A communication system consists of n components, each of which will, independently, function with probability p. The total system will be able to operate effectively if at least one-half of its components function. For what values of p is a 5-component system more likely to operate effectively than a 3-component system?

Solution

The number of functioning components is a binomial RV with parameters (n, p). The 5-component system is better if

$$P\{\text{effective 5-comp. sys.}\} > P\{\text{effective 3-comp. sys.}\}$$

$$\binom{5}{3} p^3 (1-p)^2 + \binom{5}{4} p^4 (1-p) + p^5 > \binom{3}{2} p^2 (1-p) + p^3$$

which reduces to

$$3(p-1)^2(2p-1) > 0 \qquad \Rightarrow p > \tfrac{1}{2}$$

Binomial Distribution

$$\begin{aligned} E[X^k] &= \sum_{i=0}^{n} i^k \binom{n}{i} p^i q^{n-i} \\ &= \sum_{i=1}^{n} i^k \binom{n}{i} p^i q^{n-i} \\ &= \sum_{i=1}^{n} i^k \frac{n!}{i!(n-i)!} p^i q^{n-i} \\ &= np \sum_{i=1}^{n} i^{k-1} \frac{(n-1)!}{(i-1)!(n-i)!} p^{i-1} q^{n-i} \\ &= np \sum_{i=1}^{n} i^{k-1} \binom{n-1}{i-1} p^{i-1} q^{n-i} \\ &= np \sum_{j=0}^{n-1} (j+1)^{k-1} \binom{n-1}{j} p^j q^{n-1-j} \\ &= npE\left[(Y+1)^{k-1}\right] \end{aligned}$$

$Y \sim \mathrm{Binomial}(n-1, p)$

$$\begin{aligned} E[X] &= npE\left[(Y+1)^0\right] \\ &= npE[1] \\ &= np \\ E[X^2] &= npE[Y+1] \\ &= np\left((n-1)p + 1\right) \\ &= n^2p^2 - np^2 + np \\ \mathrm{Var}(X) &= E[X^2] - \left(E[X]\right)^2 \\ &= np - np^2 \\ &= np(1-p) \\ &= npq \end{aligned}$$

Binomial Distribution

Binomial RV as a sum of n independent identical Bernoulli RVs

$X = X_1 + X_2 + \cdots + X_n \qquad (X \sim \mathrm{Binomial}(n,p), X_i \sim \mathrm{Bernoulli}(p), \text{indep.} X_i\text{'s})$

$$\begin{aligned} E[X] &= E[X_1] + E[X_2] + \cdots + E[X_n] \\ &= p + p + \cdots + p \\ &= np \end{aligned}$$

$$\begin{aligned} \mathrm{Var}(X) &= \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \cdots + \mathrm{Var}(X_n) \\ &= pq + pq + \cdots + pq \\ &= npq \end{aligned}$$

Computing the binomial distribution function

Suppose that $X \sim \mathrm{Binomial}(n, p)$. The key to computing its distribution function

$$P\{X \le i\} = \sum_{k=0}^{i} \binom{n}{i} p^i (1-p)^{n-i} \qquad i = 0, 1, \cdots, n$$

is to start with $P\{X = 0\}$ and then to compute $P\{X = k+1\}$ from $P\{X = k\}$ using the relationship

$$P\{X = k+1\} = \frac{p}{1-p} \frac{n-k}{k+1} P\{X = k\}$$

Poisson Distribution

The Poisson Random Variable

A random variable X that takes on one of the values $0, 1, 2, \cdots$ is said to be a *Poisson random variable* with parameter λ if, for some $\lambda > 0$,

$$p(i) = P\{X = i\} = e^{-\lambda}\frac{\lambda^i}{i!} \qquad i = 0, 1, 2, \cdots$$

$$\sum_{i=0}^{\infty} p(i) = \sum_{i=0}^{\infty} e^{-\lambda}\frac{\lambda^i}{i!} = e^{-\lambda}\sum_{i=0}^{\infty}\frac{\lambda^i}{i!} = e^{-\lambda}e^{\lambda} = 1$$

Poisson Distribution

Poisson RV as an approximation to binomial RV

The Poisson RV with parameter $\lambda = np$ may be used as an approximation for a binomial RV with parameters (n, p) when n is large and p is small enough so that np is of moderate size.

$$P\{X = i\} = \binom{n}{i} p^i (1-p)^{n-i} = \frac{n!}{(n-i)!i!}\left(\frac{\lambda}{n}\right)^i\left(1-\frac{\lambda}{n}\right)^{n-i} = \frac{n(n-1)\cdots(n-i+1)}{n^i}\times\frac{\lambda^i}{i!}\times\frac{(1-\lambda/n)^n}{(1-\lambda/n)^i}$$

for large n, small p, moderate np

$$\approx 1 \times \frac{\lambda^i}{i!} \times \frac{e^{-\lambda}}{1} = e^{-\lambda}\frac{\lambda^i}{i!}$$

Poisson Distribution

Examples of RVs that generally obey the Poisson probability law

- The number of misprints on a page (or a group of pages) of a book
- The number of wrong telephone numbers that are dialed in a day
- The number of customers entering a post office on a given day
- The number of vacancies occurring during a year in the federal judicial system
- The number of α-particles discharged in a fixed period of time from some radioactive material

The value λ will usually be determined empirically

Poisson Distribution

Example 1

Suppose that the number of typographical errors on a single page of our textbook has a Poisson distribution with parameter $\lambda = 1/2$. Calculate the probability that there is at least one error on page 76.

Letting X denote the number of errors on this page

$$P\{X \geq 1\} = 1 - P\{X = 0\} = 1 - e^{-1/2} \approx .393$$

Example 2

Suppose that the probability that an item produced by a certain machine will be defective is .1. Find the probability that a sample of 10 items will contain at most 1 defective item.

$$\binom{10}{0}(.1)^0(.9)^10 + \binom{10}{1}(.1)^1(.9)^9 = .7361$$

$$e^{-1} + e^{-1} \approx .7358$$

Poisson Distribution

Example 3

Consider an experiment that consists of counting the number of α particles given off in a 1-second interval by 1 gram of radioactive material. If we know from past experience that, on the average, 3.2 such α particles are given off, what is a good approximation to the probability that no more than 2 α particles will appear?

- Think of the gram of radioactive material as consisting of a large number n of atoms, each of which has probability of $3.2/n$ of disintegrating and sending off an α particle during the second considered
- The number of α particles given off will be a Poisson random variable with parameter $\lambda = 3.2$

$$P\{X \leq 2\} = e^{-3.2} + 3.2e^{-3.2} + \frac{3.2^2}{2}e^{-3.2} \approx .3799$$

Poisson Distribution

Binomial Approximation

$$\lambda = np$$
$$q = 1$$
$$E[X] = np = \lambda$$
$$\mathrm{Var}(X) = npq = \lambda$$

$$E[X] = \sum_{i=0}^{\infty}\frac{ie^{-\lambda}\lambda^i}{i!} = \lambda\sum_{i=1}^{\infty}\frac{e^{-\lambda}\lambda^{i-1}}{(i-1)!} = \lambda e^{-\lambda}\sum_{j=0}^{\infty}\frac{\lambda^j}{j!} = \lambda e^{-\lambda}e^{\lambda} = \lambda$$

$$E[X^2] = \sum_{i=0}^{\infty}\frac{i^2e^{-\lambda}\lambda^i}{i!} = \lambda\sum_{i=1}^{\infty}\frac{ie^{-\lambda}\lambda^{i-1}}{(i-1)!} = \lambda\sum_{j=0}^{\infty}\frac{(j+1)e^{-\lambda}\lambda^j}{j!} = \lambda\left(\sum_{j=0}^{\infty}\frac{je^{-\lambda}\lambda^j}{j!} + \sum_{j=0}^{\infty}\frac{e^{-\lambda}\lambda^j}{j!}\right) = \lambda(\lambda+1)$$

$$\mathrm{Var}(X) = E[X^2] - \left(E[X]\right)^2 = \lambda$$

Poisson Distribution

Computing the Poisson distribution function

Suppose that $X \sim \text{Poisson}(\lambda)$. The key to computing its distribution function

$$P\{X \leq i\} = \sum_{k=0}^{i} \frac{e^{-\lambda}\lambda^k}{k!} \qquad i = 0, 1, 2, \cdots$$

is to start with $P\{X = 0\}$ and then to compute $P\{X = k+1\}$ from $P\{X = k\}$ using the relationship

$$P\{X = k+1\} = \frac{\lambda}{k+1} P\{X = k\}$$

Geometric Distribution

The Geometric Random Variable

Suppose that independent trials, each having a probability p, $0 < p < 1$, of being a success, are performed until a success occurs. If we let X equal the number of trials required, then

$$P\{X = i\} = q^{i-1}p \qquad q = 1 - p, \qquad i = 1, 2, \cdots$$

$$\begin{aligned}
\sum_{i=0}^{\infty} P\{X = i\} &= \sum_{i=1}^{\infty} q^{i-1}p \\
&= \frac{p}{1-q} \\
&= \frac{p}{p} \\
&= 1
\end{aligned}$$

$$\begin{aligned}
P\{X \leq i\} &= \sum_{k=1}^{i} q^{k-1}p \\
&= \frac{p - q^i p}{1-q} \\
&= \frac{p(1-q^i)}{p} \\
&= 1 - q^i
\end{aligned}$$

Geometric Distribution

Example

An urn contains N white and M black balls. Balls are randomly selected, one at a time, until a black one is obtained. If we assume that each selected ball is replaced before the next one is drawn, what is the probability that

1. exactly n draws are needed?
2. at least n draws are needed?

- Let X denote the number of draws needed to select a black ball
- $X \sim \text{Geometric}(p), \qquad p = M/(M+N), \qquad q = N/(M+N)$

1. $P\{X = n\} = q^{n-1}p = \left(\frac{N}{M+N}\right)^{n-1} \frac{M}{M+N} = \frac{MN^{n-1}}{(M+N)^n}$
2. $P\{X \geq n\} = \sum_{k=n}^{\infty} q^{k-1}p = q^{n-1} = \left(\frac{N}{M+N}\right)^{n-1}$

Geometric Distribution

$$\begin{aligned}
E[X] &= \sum_{i=1}^{\infty} iq^{i-1}p \\
&= \sum_{i=1}^{\infty} (i-1+1)q^{i-1}p \\
&= \sum_{j=0}^{\infty} (j+1)q^j p \\
&= \sum_{j=0}^{\infty} jq^j p + \sum_{j=0}^{\infty} q^j p \\
&= q\sum_{j=1}^{\infty} jq^{j-1}p + \sum_{i=1}^{\infty} q^{i-1}p \\
&= qE[X] + 1 \\
&= 1/(1-q) \\
E[X] &= 1/p
\end{aligned}$$

$$\begin{aligned}
E[X^2] &= \sum_{i=1}^{\infty} i^2 q^{i-1}p \\
&= \sum_{i=1}^{\infty} (i-1+1)^2 q^{i-1}p \\
&= \sum_{j=0}^{\infty} (j+1)^2 q^j p \\
&= \sum_{j=0}^{\infty} j^2 q^j p + 2\sum_{j=0}^{\infty} jq^j p + \sum_{j=0}^{\infty} q^j p \\
&= q\sum_{j=1}^{\infty} j^2 q^{j-1}p + 2q\sum_{j=1}^{\infty} jq^{j-1}p + \sum_{i=1}^{\infty} q^{i-1}p \\
&= qE[X^2] + 2qE[X] + 1 \\
&= (2q/p + 1)/(1-q) \\
&= (q+1)/p^2
\end{aligned}$$

$$\text{Var}(X) = E[X^2] - (E[X])^2 \qquad = q/p^2$$