

3. Random Variables

Let (Ω, F, P) be a probability model for an experiment, and X a function that maps every $\xi \in \Omega$, to a unique point $x \in R$, the set of real numbers. Since the outcome ξ is not certain, so is the value $X(\xi) = x$. Thus if B is some subset of R , we may want to determine the probability of “ $X(\xi) \in B$ ”. To determine this probability, we can look at the set $A = X^{-1}(B) \in \Omega$ that contains all $\xi \in \Omega$ that maps into B under the function X .

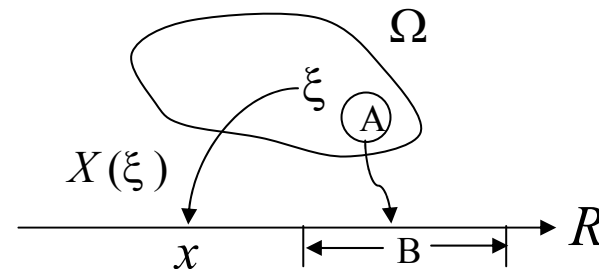


Fig. 3.1

Obviously, if the set $A = X^{-1}(B)$ also belongs to the associated field F , then it is an event and the probability of A is well defined; in that case we can say

$$\text{Probability of the event " } X(\xi) \in B \text{ " = } P(X^{-1}(B)). \quad (3-1)$$

However, $X^{-1}(B)$ may not always belong to F for all B , thus creating difficulties. The notion of random variable (r.v) makes sure that the inverse mapping always results in an event so that we are able to determine the probability for any $B \in R$.

Random Variable (r.v): A finite single valued function $X(\cdot)$ that maps the set of all experimental outcomes Ω into the set of real numbers R is said to be a r.v, if the set $\{\xi \mid X(\xi) \leq x\}$ is an event ($\in F$) for every x in R .

Alternatively X is said to be a r.v, if $X^{-1}(B) \in F$ where B represents semi-definite intervals of the form $\{-\infty < x \leq a\}$ and all other sets that can be constructed from these sets by performing the set operations of union, intersection and negation any number of times. The Borel collection B of such subsets of R is the smallest σ -field of subsets of R that includes all semi-infinite intervals of the above form. Thus if X is a r.v, then

$$\{\xi \mid X(\xi) \leq x\} = \{X \leq x\} \quad (3-2)$$

is an event for every x . What about $\{a < X \leq b\}$, $\{X = a\}$? Are they also events? In fact with $b > a$ since $\{X \leq a\}$ and $\{X \leq b\}$ are events, $\{X \leq a\}^c = \{X > a\}$ is an event and hence $\{X > a\} \cap \{X \leq b\} = \{a < X \leq b\}$ is also an event.

Thus, $\left\{ a - \frac{1}{n} < X \leq a \right\}$ is an event for every n .

Consequently

$$\bigcap_{n=1}^{\infty} \left\{ a - \frac{1}{n} < X \leq a \right\} = \{ X = a \} \quad (3-3)$$

is also an event. All events have well defined probability. Thus the probability of the event $\{ \xi \mid X(\xi) \leq x \}$ must depend on x . Denote

$$P \{ \xi \mid X(\xi) \leq x \} = F_X(x) \geq 0. \quad (3-4)$$

The role of the subscript X in (3-4) is only to identify the actual r.v. $F_X(x)$ is said to be the Probability Distribution Function (PDF) associated with the r.v. X .

Distribution Function: Note that a distribution function $g(x)$ is nondecreasing, right-continuous and satisfies

$$g(+\infty) = 1, \quad g(-\infty) = 0, \quad (3-5)$$

i.e., if $g(x)$ is a distribution function, then

(i) $g(+\infty) = 1, \quad g(-\infty) = 0,$

(ii) if $x_1 < x_2$, then $g(x_1) \leq g(x_2)$, (3-6)

and

(iii) $g(x^+) = g(x)$, for all x .

We need to show that $F_X(x)$ defined in (3-4) satisfies all properties in (3-6). In fact, for any r.v X ,

$$(i) \quad F_X(+\infty) = P\{\xi \mid X(\xi) \leq +\infty\} = P(\Omega) = 1 \quad (3-7)$$

$$\text{and} \quad F_X(-\infty) = P\{\xi \mid X(\xi) \leq -\infty\} = P(\phi) = 0. \quad (3-8)$$

(ii) If $x_1 < x_2$, then the subset $(-\infty, x_1) \subset (-\infty, x_2)$.

Consequently the event $\{\xi \mid X(\xi) \leq x_1\} \subset \{\xi \mid X(\xi) \leq x_2\}$, since $X(\xi) \leq x_1$ implies $X(\xi) \leq x_2$. As a result

$$F_X(x_1) \triangleq P(X(\xi) \leq x_1) \leq P(X(\xi) \leq x_2) \triangleq F_X(x_2), \quad (3-9)$$

implying that the probability distribution function is nonnegative and monotone nondecreasing.

(iii) Let $x < x_n < x_{n-1} < \cdots < x_2 < x_1$, and consider the event

$$A_k = \{\xi \mid x < X(\xi) \leq x_k\} \quad (3-10)$$

since

$$\{x < X(\xi) \leq x_k\} \cup \{X(\xi) \leq x\} = \{X(\xi) \leq x_k\}, \quad (3-11) \quad 6$$

using mutually exclusive property of events we get

$$P(A_k) = P(x < X(\xi) \leq x_k) = F_X(x_k) - F_X(x). \quad (3-12)$$

But $\cdots A_{k+1} \subset A_k \subset A_{k-1} \cdots$, and hence

$$\lim_{k \rightarrow \infty} A_k = \bigcap_{k=1}^{\infty} A_k = \phi \quad \text{and hence} \quad \lim_{k \rightarrow \infty} P(A_k) = 0. \quad (3-13)$$

Thus

$$\lim_{k \rightarrow \infty} P(A_k) = \lim_{k \rightarrow \infty} F_X(x_k) - F_X(x) = 0.$$

But $\lim_{k \rightarrow \infty} x_k = x^+$, the right limit of x , and hence

$$F_X(x^+) = F_X(x), \quad (3-14)$$

i.e., $F_X(x)$ is right-continuous, justifying all properties of a distribution function.

Additional Properties of a PDF

(iv) If $F_X(x_0) = 0$ for some x_0 , then $F_X(x) = 0, x \leq x_0$. (3-15)

This follows, since $F_X(x_0) = P(X(\xi) \leq x_0) = 0$ implies $\{X(\xi) \leq x_0\}$ is the null set, and for any $x \leq x_0$, $\{X(\xi) \leq x\}$ will be a subset of the null set.

(v) $P\{X(\xi) > x\} = 1 - F_X(x)$. (3-16)

We have $\{X(\xi) \leq x\} \cup \{X(\xi) > x\} = \Omega$, and since the two events are mutually exclusive, (16) follows.

(vi) $P\{x_1 < X(\xi) \leq x_2\} = F_X(x_2) - F_X(x_1), x_2 > x_1$. (3-17)

The events $\{X(\xi) \leq x_1\}$ and $\{x_1 < X(\xi) \leq x_2\}$ are mutually exclusive and their union represents the event $\{X(\xi) \leq x_2\}$.

$$(vii) \quad P(X(\xi) = x) = F_X(x) - F_X(x^-). \quad (3-18)$$

Let $x_1 = x - \varepsilon$, $\varepsilon > 0$, and $x_2 = x$. From (3-17)

$$\lim_{\varepsilon \rightarrow 0} P\{x - \varepsilon < X(\xi) \leq x\} = F_X(x) - \lim_{\varepsilon \rightarrow 0} F_X(x - \varepsilon), \quad (3-19)$$

or

$$P\{X(\xi) = x\} = F_X(x) - F_X(x^-). \quad (3-20)$$

According to (3-14), $F_X(x_0^+)$, the limit of $F_X(x)$ as $x \rightarrow x_0$ from the right always exists and equals $F_X(x_0)$. However the left limit value $F_X(x_0^-)$ need not equal $F_X(x_0)$. Thus $F_X(x)$ need not be continuous from the left. At a discontinuity point of the distribution, the left and right limits are different, and from (3-20)

$$P\{X(\xi) = x_0\} = F_X(x_0) - F_X(x_0^-) > 0. \quad (3-21) \quad 9$$

Thus the only discontinuities of a distribution function $F_X(x)$ are of the jump type, and occur at points x_0 where (3-21) is satisfied. These points can always be enumerated as a sequence, and moreover they are at most countable in number.

Example 3.1: X is a r.v such that $X(\xi) = c, \xi \in \Omega$. Find $F_X(x)$.

Solution: For $x < c$, $\{X(\xi) \leq x\} = \{\phi\}$, so that $F_X(x) = 0$, and for $x > c$, $\{X(\xi) \leq x\} = \Omega$, so that $F_X(x) = 1$. (Fig.3.2)

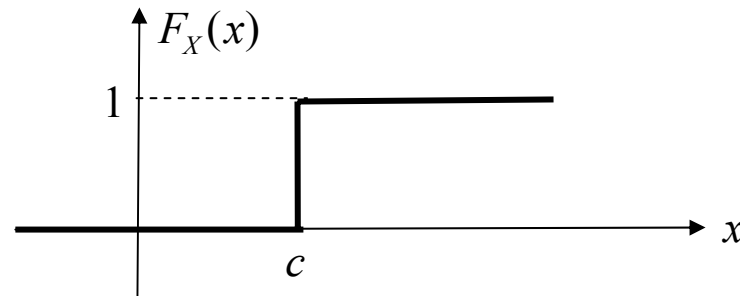


Fig. 3.2

Example 3.2: Toss a coin. $\Omega = \{H, T\}$. Suppose the r.v X is such that $X(T) = 0, X(H) = 1$. Find $F_X(x)$.

Solution: For $x < 0, \{ X(\xi) \leq x \} = \{ \phi \}$, so that $F_X(x) = 0$.

$0 \leq x < 1, \{ X(\xi) \leq x \} = \{ T \}$, so that $F_X(x) = P\{ T \} = 1 - p$,

$x \geq 1, \{ X(\xi) \leq x \} = \{ H, T \} = \Omega$, so that $F_X(x) = 1$. (Fig. 3.3)

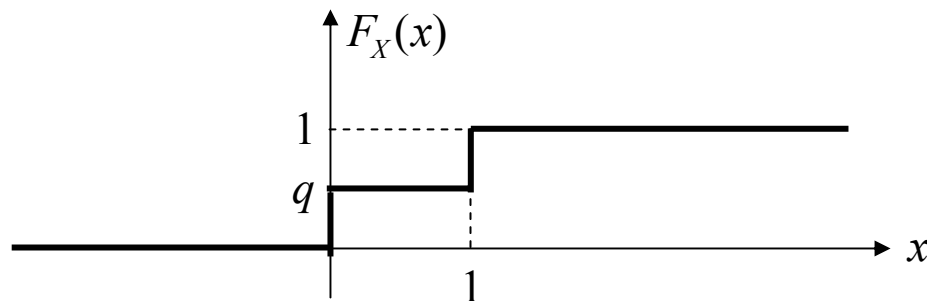


Fig.3.3

• X is said to be a continuous-type r.v if its distribution function $F_X(x)$ is continuous. In that case $F_X(x^-) = F_X(x)$ for all x , and from (3-21) we get $P\{X = x\} = 0$.

• If $F_X(x)$ is constant except for a finite number of jump discontinuities (piece-wise constant; step-type), then X is said to be a discrete-type r.v. If x_i is such a discontinuity point, then from (3-21)

$$p_i = P\{X = x_i\} = F_X(x_i) - F_X(x_i^-). \quad (3-22)$$

From Fig.3.2, at a point of discontinuity we get

$$P\{X = c\} = F_X(c) - F_X(c^-) = 1 - 0 = 1.$$

and from Fig.3.3,

$$P\{X = 0\} = F_X(0) - F_X(0^-) = q - 0 = q.$$

Example:3.3 A fair coin is tossed twice, and let the r.v X represent the number of heads. Find $F_X(x)$.

Solution: In this case $\Omega = \{HH, HT, TH, TT\}$, and

$$X(HH) = 2, X(HT) = 1, X(TH) = 1, X(TT) = 0.$$

$$x < 0, \{X(\xi) \leq x\} = \phi \Rightarrow F_X(x) = 0,$$

$$0 \leq x < 1, \{X(\xi) \leq x\} = \{TT\} \Rightarrow F_X(x) = P\{TT\} = P(T)P(T) = \frac{1}{4},$$

$$1 \leq x < 2, \{X(\xi) \leq x\} = \{TT, HT, TH\} \Rightarrow F_X(x) = P\{TT, HT, TH\} = \frac{3}{4},$$

$$x \geq 2, \{X(\xi) \leq x\} = \Omega \Rightarrow F_X(x) = 1. \text{ (Fig. 3.4)}$$

From Fig.3.4, $P\{X = 1\} = F_X(1) - F_X(1^-) = 3/4 - 1/4 = 1/2$.

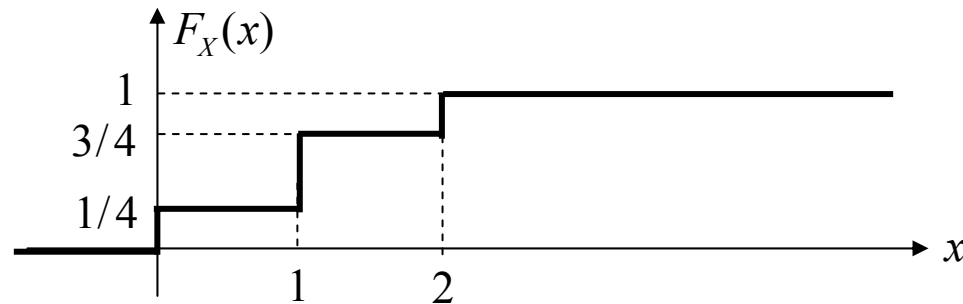


Fig. 3.4

Probability density function (p.d.f)

The derivative of the distribution function $F_X(x)$ is called the probability density function $f_X(x)$ of the r.v X . Thus

$$f_X(x) \triangleq \frac{dF_X(x)}{dx}. \quad (3-23)$$

Since

$$\frac{dF_X(x)}{dx} = \lim_{\Delta x \rightarrow 0} \frac{F_X(x + \Delta x) - F_X(x)}{\Delta x} \geq 0, \quad (3-24)$$

from the monotone-nondecreasing nature of $F_X(x)$,

it follows that $f_X(x) \geq 0$ for all x . $f_X(x)$ will be a continuous function, if X is a continuous type r.v.

However, if X is a discrete type r.v as in (3-22), then its p.d.f has the general form (Fig. 3.5)

$$f_X(x) = \sum_i p_i \delta(x - x_i), \quad (3-25)$$

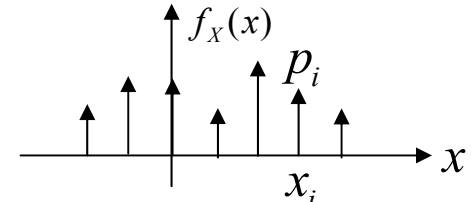


Fig. 3.5

where x_i represent the jump-discontinuity points in $F_X(x)$. As Fig. 3.5 shows $f_X(x)$ represents a collection of positive discrete masses, and it is known as the probability mass function (p.m.f) in the discrete case. From (3-23), we also obtain by integration

$$F_X(x) = \int_{-\infty}^x f_X(u) du. \quad (3-26)$$

Since $F_X(+\infty) = 1$, (3-26) yields

$$\int_{-\infty}^{+\infty} f_X(x) dx = 1, \quad (3-27)$$

which justifies its name as the density function. Further, from (3-26), we also get (Fig. 3.6b)

$$P \{ x_1 < X (\xi) \leq x_2 \} = F_X(x_2) - F_X(x_1) = \int_{x_1}^{x_2} f_X(x) dx . \quad (3-28)$$

Thus the area under $f_X(x)$ in the interval (x_1, x_2) represents the probability in (3-28).

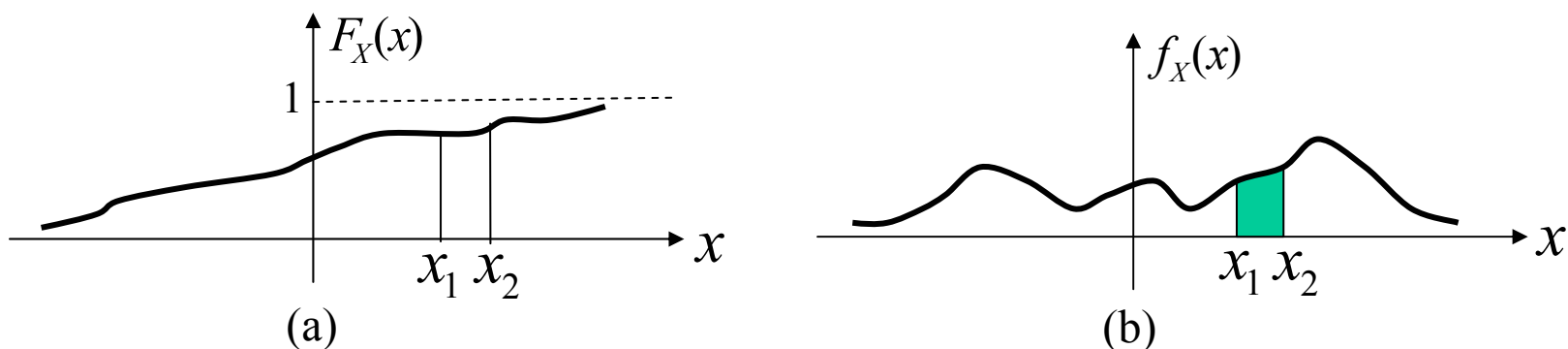


Fig. 3.6

Often, r.v.s are referred by their specific density functions - both in the continuous and discrete cases - and in what follows we shall list a number of them in each category.

Continuous-type random variables

1. Normal (Gaussian): X is said to be normal or Gaussian r.v, if

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2 / 2\sigma^2}. \quad (3-29)$$

This is a bell shaped curve, symmetric around the parameter μ , and its distribution function is given by

$$F_X(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(y-\mu)^2 / 2\sigma^2} dy \triangleq G\left(\frac{x-\mu}{\sigma}\right), \quad (3-30)$$

where $G(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy$ is often tabulated. Since $f_X(x)$ depends on two parameters μ and σ^2 , the notation $X \sim N(\mu, \sigma^2)$ will be used to represent (3-29).

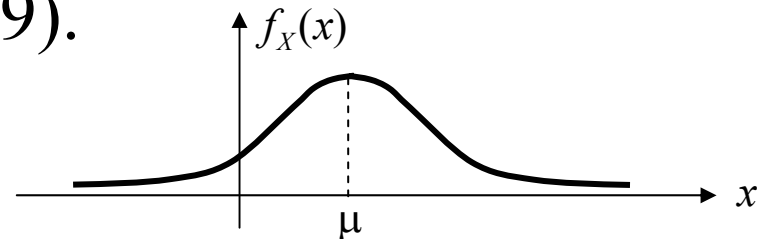


Fig. 3.7

2. Uniform: $X \sim U(a,b)$, $a < b$, if (Fig. 3.8)

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b, \\ 0, & \text{otherwise.} \end{cases} \quad (3.31)$$

3. Exponential: $X \sim \varepsilon(\lambda)$ if (Fig. 3.9)

$$f_X(x) = \begin{cases} \frac{1}{\lambda} e^{-x/\lambda}, & x \geq 0, \\ 0, & \text{otherwise.} \end{cases} \quad (3-32)$$

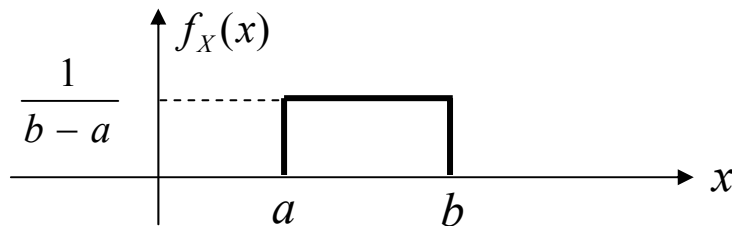


Fig. 3.8

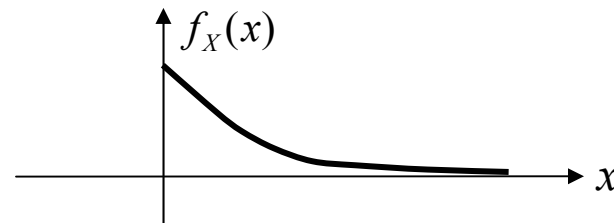


Fig. 3.9

4. Gamma: $X \sim G(\alpha, \beta)$ if $(\alpha > 0, \beta > 0)$ (Fig. 3.10)

$$f_X(x) = \begin{cases} \frac{x^{\alpha-1}}{\Gamma(\alpha)\beta^\alpha} e^{-x/\beta}, & x \geq 0, \\ 0, & \text{otherwise.} \end{cases} \quad (3-33)$$

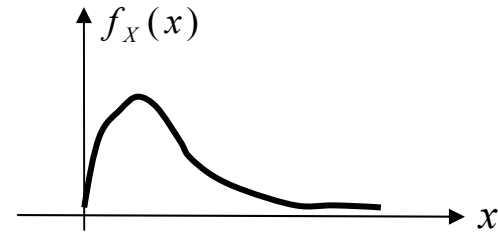


Fig. 3.10

If $\alpha = n$ an integer $\Gamma(n) = (n - 1)!$.

5. Beta: $X \sim \beta(a, b)$ if $(a > 0, b > 0)$ (Fig. 3.11)

$$f_X(x) = \begin{cases} \frac{1}{\beta(a, b)} x^{a-1} (1-x)^{b-1}, & 0 < x < 1, \\ 0, & \text{otherwise.} \end{cases} \quad (3-34)$$

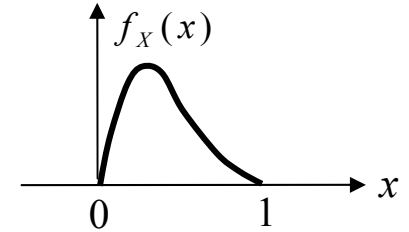


Fig. 3.11

where the Beta function $\beta(a, b)$ is defined as

$$\beta(a, b) = \int_0^1 u^{a-1} (1-u)^{b-1} du. \quad (3-35)$$

6. Chi-Square: $X \sim \chi^2(n)$, if (Fig. 3.12)

$$f_X(x) = \begin{cases} \frac{1}{2^{n/2} \Gamma(n/2)} x^{n/2-1} e^{-x/2}, & x \geq 0, \\ 0, & \text{otherwise.} \end{cases} \quad (3-36)$$

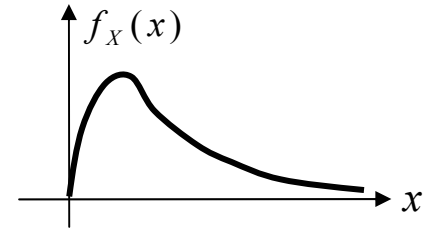


Fig. 3.12

Note that $\chi^2(n)$ is the same as Gamma $(n/2, 2)$.

7. Rayleigh: $X \sim R(\sigma^2)$, if (Fig. 3.13)

$$f_X(x) = \begin{cases} \frac{x}{\sigma^2} e^{-x^2/2\sigma^2}, & x \geq 0, \\ 0, & \text{otherwise.} \end{cases} \quad (3-37)$$

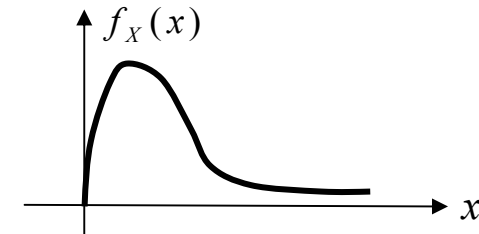


Fig. 3.13

8. Nakagami – m distribution:

$$f_X(x) = \begin{cases} \frac{2}{\Gamma(m)} \left(\frac{m}{\Omega}\right)^m x^{2m-1} e^{-mx^2/\Omega}, & x \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (3-38)$$

9. Cauchy: $X \sim C(\alpha, \mu)$, if (Fig. 3.14)

$$f_X(x) = \frac{\alpha / \pi}{\alpha^2 + (x - \mu)^2}, \quad -\infty < x < +\infty. \quad (3-39)$$

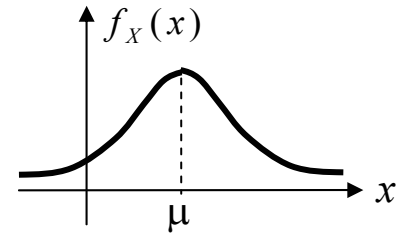


Fig. 3.14

10. Laplace: (Fig. 3.15)

$$f_X(x) = \frac{1}{2\lambda} e^{-|x|/\lambda}, \quad -\infty < x < +\infty. \quad (3-40)$$

11. Student's t -distribution with n degrees of freedom (Fig 3.16)

$$f_T(t) = \frac{\Gamma((n+1)/2)}{\sqrt{\pi n} \Gamma(n/2)} \left(1 + \frac{t^2}{n}\right)^{-(n+1)/2}, \quad -\infty < t < +\infty. \quad (3-41)$$

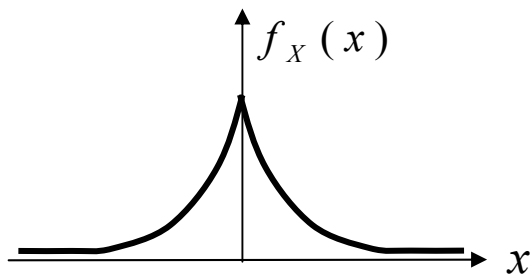


Fig. 3.15

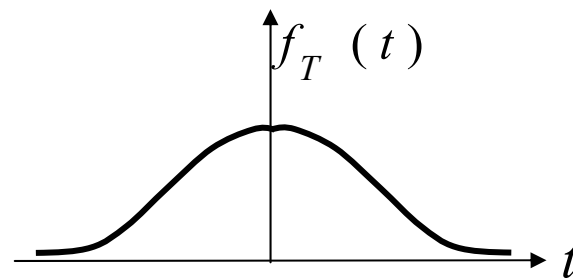


Fig. 3.16

12. Fisher's F-distribution

$$f_z(z) = \begin{cases} \frac{\Gamma\{(m+n)/2\} m^{m/2} n^{n/2}}{\Gamma(m/2)\Gamma(n/2)} \frac{z^{m/2-1}}{(n+mz)^{(m+n)/2}}, & z \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (3-42)$$

Discrete-type random variables

1. Bernoulli: X takes the values (0,1), and

$$P(X = 0) = q, \quad P(X = 1) = p. \quad (3-43)$$

2. Binomial: $X \sim B(n, p)$, if (Fig. 3.17)

$$P(X = k) = \binom{n}{k} p^k q^{n-k}, \quad k = 0, 1, 2, \dots, n. \quad (3-44)$$

3. Poisson: $X \sim P(\lambda)$, if (Fig. 3.18)

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \dots, \infty. \quad (3-45)$$

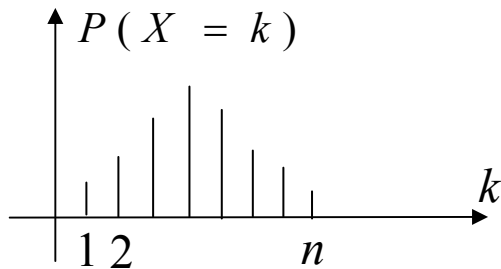


Fig. 3.17

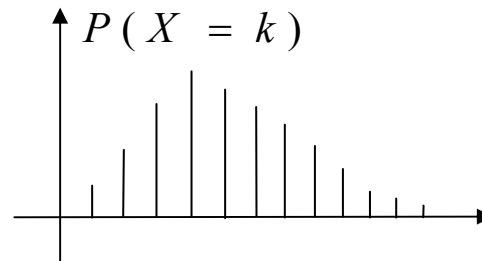


Fig. 3.18

4. Hypergeometric:

$$P(X = k) = \frac{\binom{m}{k} \binom{N-m}{n-k}}{\binom{N}{n}}, \quad \max(0, m + n - N) \leq k \leq \min(m, n) \quad (3-46)$$

5. Geometric: $X \sim g(p)$ if

$$P(X = k) = pq^k, \quad k = 0, 1, 2, \dots, \infty, \quad q = 1 - p. \quad (3-47)$$

6. Negative Binomial: $X \sim NB(r, p)$, if

$$P(X = k) = \binom{k-1}{r-1} p^r q^{k-r}, \quad k = r, r+1, \dots. \quad (3-48)$$

7. Discrete-Uniform:

$$P(X = k) = \frac{1}{N}, \quad k = 1, 2, \dots, N. \quad (3-49)$$

We conclude this lecture with a general distribution due

to Polya that includes both binomial and hypergeometric as special cases.

Polya's distribution: A box contains a white balls and b black balls. A ball is drawn at random, and it is replaced along with c balls of the same color. If X represents the number of white balls drawn in n such draws, $X = 0, 1, 2, \dots, n$, find the probability mass function of X .

Solution: Consider the specific sequence of draws where k white balls are first drawn, followed by $n - k$ black balls. The probability of drawing k successive white balls is given by

$$p_w = \frac{a}{a+b} \frac{a+c}{a+b+c} \frac{a+2c}{a+b+2c} \dots \frac{a+(k-1)c}{a+b+(k-1)c} \quad (3-50)$$

Similarly the probability of drawing k white balls

followed by $n - k$ black balls is given by

$$\begin{aligned}
 p_k &= p_w \frac{b}{a+b+kc} \frac{b+c}{a+b+(k+1)c} \cdots \frac{b+(n-k-1)c}{a+b+(n-1)c} \\
 &= \prod_{i=0}^{k-1} \frac{a+ic}{a+b+ic} \prod_{j=0}^{n-k-1} \frac{b+jc}{a+b+(j+k)c}.
 \end{aligned} \tag{3-51}$$

Interestingly, p_k in (3-51) also represents the probability of drawing k white balls and $(n - k)$ black balls in *any other specific order* (i.e., The same set of numerator and denominator terms in (3-51) contribute to *all other* sequences as well.) But there are $\binom{n}{k}$ such distinct mutually exclusive sequences and summing over all of them, we obtain the Polya distribution (probability of getting k white balls in n draws) to be

$$P(X = k) = \binom{n}{k} p_k = \binom{n}{k} \prod_{i=0}^{k-1} \frac{a+ic}{a+b+ic} \prod_{j=0}^{n-k-1} \frac{b+jc}{a+b+(j+k)c}, \quad k = 0, 1, 2, \dots, n. \tag{3-52}$$

PILLAI

Both binomial distribution as well as the hypergeometric distribution are special cases of (3-52).

For example if draws are done with replacement, then $c = 0$ and (3-52) simplifies to the binomial distribution

$$P(X = k) = \binom{n}{k} p^k q^{n-k}, \quad k = 0, 1, 2, \dots, n \quad (3-53)$$

where

$$p = \frac{a}{a+b}, \quad q = \frac{b}{a+b} = 1 - p.$$

Similarly if the draws are conducted without replacement, Then $c = -1$ in (3-52), and it gives

$$P(X = k) = \binom{n}{k} \frac{a(a-1)(a-2)\cdots(a-k+1)}{(a+b)(a+b-1)\cdots(a+b-k+1)} \frac{b(b-1)\cdots(b-n+k+1)}{(a+b-k)\cdots(a+b-n+1)}$$

$$P(X = k) = \frac{n!}{k!(n-k)!} \frac{a!(a+b-k)!}{(a-k)!(a+b)!} \frac{b!(a+b-n)!}{(b-n+k)!(a+b-k)!} = \frac{\binom{a}{k} \binom{b}{n-k}}{\binom{a+b}{n}} \quad (3-54)$$

which represents the hypergeometric distribution. Finally $c = +1$ gives (replacements are doubled)

$$P(X = k) = \binom{n}{k} \frac{(a+k-1)! (a+b+1)!}{(a-1)! (a+b+k-1)!} \frac{(b+n-k-1)! (a+b+k-1)!}{(b-1)! (a+b+n-1)!} \\ = \frac{\binom{a+k-1}{k} \binom{b+n-k-1}{n-k}}{\binom{a+b+n-1}{n}}. \quad (3-55)$$

we shall refer to (3-55) as Polya's +1 distribution. the general Polya distribution in (3-52) has been used to study the spread of contagious diseases (epidemic modeling).