船舶与海洋工程系列
CHUANBO YU HAIYANG GONGCHENG XILIE

船舶强度的概率方法

杨代盛 桑国光 李维扬 戴仰山 ◎编 著

哈尔滨工程大学出版社
Harbin Engineering University Press

船舶与海洋工程系列

CHUANBO YU HAIYANG GONGCHENG XILIE

船舶强度的概率方法

杨代盛 桑国光 李维扬 戴仰山◎编 著

哈尔滨工程大学出版社
Harbin Engineering University Press

内容简介

本书共分四章。第一章介绍概率论及平稳随机过程的基本知识；第二章介绍按切片理论计算船舶波浪载荷的方法；第三章介绍船舶在波浪中砰击和弹振发生的机理及有关特性估算；第四章介绍船舶总纵强度的结构可靠性分析。本书的特点是着重物理概念的阐述，介绍目前比较成熟的按概率理论研究船舶强度的有关方法。

本书可作为高等学校船舶与海洋工程结构力学专业的硕士研究生教材，亦可供从事船舶结构强度教学的教师、科研工作者及工程技术人员参考。

ISBN 978-7-81007-491-9

出版发行 哈尔滨工程大学出版社
社　　址 哈尔滨市南岗区东大直街124号
邮政编码 150001
发行电话 0451-82519328
传　　真 0451-82519699
经　　销 新华书店
印　　刷 黑龙江省教育厅印刷厂
开　　本 787mm×1 092mm　1/16
印　　张 11
字　　数 254千字
版　　次 2007年10月修订
印　　次 2007年10月第2次印刷
定　　价 22.00元
http://press.hrbeu.edu.cn
E-mail:heupress@hrbeu.edu.cn

前　言

船舶强度的任务是研究船体结构在外力作用下抵抗破坏的能力和规律。把船作为一个整体来研究它的强度,称之为船体总强度问题。由于船体细而长,主要是纵向弯曲变形,所以船体总强度主要是指船体总纵弯曲时的强度。研究组成船体的部分结构、节点及其组成构件的强度,则称为船体局部强度。

船舶在波浪中总纵弯曲时的外力,长期以来是按照将船静置于标准波浪上的静力方法来确定。在计算船体的总纵弯曲应力时,采用的亦是确定性方法。事实上,不仅作用于船体的波浪载荷是随机变化的,由于钢材冶炼及船舶建造过程中许多实际因素的影响,船体的抗弯能力也不是个确定性的量。基于此,为了合理地评断船体结构的安全性,应当采用概率统计理论来研究船舶强度。近些年来,国内外在这方面的研究有了很大的进展。

为了介绍和推广这种新方法,中国造船工程学会力学学术委员会曾于1981年夏在上海举办了“近代船舶强度理论”学习班(邀请杨代盛、桑国光、戴仰山、沈进威和于宝海等同志主讲)。之后,原哈尔滨船舶工程学院和上海交通大学相继编写并出版了《船舶概率强度》讲义,分别在教学中试用。大约在1985年左右,教材会议决定由上述两个院校共同编写这本《船舶强度的概率方法》教材,以供船舶与海洋工程结构力学专业的硕士研究生及有关人员使用。

本书共四章。第一章由李维扬编写,第二章由杨代盛编写,第三章由戴仰山编写,第四章由桑国光编写。作为当时力学学术委员会波浪载荷与动力响应学组组长的杨代盛教授,为组织本书的编写付出了艰辛的劳动。戴仰山教授负责有关出版的具体事宜,并在付印之前对文稿进行了必要的修改。

科技的发展日新月异,本书不可能对船舶强度研究中的新成果一一涉及。作为一本教材,编者是希望通过对目前比较成熟、比较实用的有关方法的介绍,着重阐明物理现象的实质。此外,在编写过程中没有过份强调全书符号、用语和风格的统一,各章保持相对的独立。

本书经中国船舶科学研究中心周国华高级工程师全面审阅,并提出了许多宝贵意见。在此,特表谢意。

遗憾的是,为我国船舶结构可靠性研究作出突出贡献的桑国光教授,因病于1990年与世长辞,没能亲眼看到本书的出版。

在出版过程中,得到了哈尔滨工程大学出版社的大力支持,在此一并表示谢意。

限于我们的水平,加之两校编者远隔两地联系不便,从编写到出版之间的时间又拖得较长,书中难免有不妥和谬误之处,恳切欢迎读者批评指正。

编　者

目　录

第 1 章

概率方法基础

符 号 说 明

$C_x(t_1,t_2)$	随机过程协方差函数
D[ ]	方差运算符号
D_x	方差
D_x^*	实验(统计)方差
E[ ]	数学期望运算符号
$F(x)$	分布函数
$f(x)$	密度函数
$g(y)$	y 的概率密度
$H(\omega)$	频率响应函数
$h(\tau)$	脉冲响应函数
I	雅可比矩阵
L	线性变换算子
$m_x(t)$	随机过程均值
m_x	均值
m_k	k 阶谱矩
m_x^*	实验(统计)均值
$N(t)$	计数过程
$N^+(\xi)$	单位时间内过阈(ξ)次数
$N^+(0)$	单位时间内过零次数
$P\{\ \}$	概率
p_i^*	在 i 区间内出现的频率
$R_x(t_1,t_2)$	随机过程相关函数、相关矩
$R_x(\tau)$	平稳过程的相关函数
$R_{xy}(t_1,t_2)$	随机过程互相关函数
$R=2D_x$	雷利分布参数
$r_x(t_1,t_2)$	随机过程标准相关函数
r_{xy}	相关系数
$S_x(\omega)$	平稳过程的单边谱

$S_x^*(\omega)$	平稳过程的双边谱
$X(t)$	随机过程
$x(t)$	随机过程的现实
$X, Y\cdots$	随机变量
$x, y\cdots$	随机变量的可能值
$\bar{y}_n$	最可能极值
$\hat{y}_n(\alpha)$	设计极值
α_k	k 阶原点矩
α_{ks}	ks 阶原点矩
α_{11}	二阶混合原点矩、相关矩
Γ	伽马函数
δ	小参数
ϵ	带宽系数
η	赖斯分布峰值参数
λ	指数分布参数
μ_k	k 阶中心矩
μ_{ks}	ks 阶中心矩
$\mu_{11} = K_{xy}$	二阶混合中心矩、协方差
$\sigma_x = \lvert D_x^{\frac{1}{2}} \rvert$	标准差
$\Phi(\ )$	拉普拉斯函数
ω_e	遭遇频率

1.1 随 机 变 量

一、随机变量、分布律与数字特征

由于人们对客观事物的了解常不能完整,从而对事物未来的发展便没有充分把握,因此就需要用随机量来对事物作量的分析。

一个可取多种可能值,但不能事先知道实验结果中具体将取哪一值的变量称为**随机变量**。根据可能值只为离散值(例如变量表示事件出现的次数),或可取给定区间内任意值(例如海浪参数、船舶摇摆幅、船体构件中应力值等),随机变量可分为离散型和连续型两类。本节中讨论的都是连续型随机变量。

为了描述随机变量,需要知道变量的全部可能值和取这些可能值的概率,这两者合起来构成随机变量的**分布律**。

今后将随机变量以大写字母 X、Y、…表示,而它们相应的可能值以小写字母 x、y、…表示。

变量 X 取值小于给定可能值 x 的概率称为 X 的**分布函数** $F(x)$,即

$$F(x) = P\{X < x\} \tag{1-1}$$

分布函数为无因次函数。它是 x 的非降函数,且 $F(-\infty)=0$, $F(\infty)=1$。注意,连续

型随机变量取某一个别值(例如 $x = a$)的概率 $P\{X = a\} = 0$。这并不是说在实验中不可能出现这个值,但其出现频率非常接近于零。所以,对连续型随机变量,只有讨论它取值在某一区间内才有意义,而且不必区分该区间是开区间或闭区间。

有了 $F(x)$,则 X 取值在$[\alpha,\beta]$内的概率为

$$P\{\alpha < X < \beta\} = F(\beta) - F(\alpha) \tag{1-2}$$

对连续型值随机变量,常用的分布律形式还有**密度函数** $f(x)$。它是分布函数的导数,即

$$f(x) = \frac{\mathrm{d}F(x)}{\mathrm{d}x} \tag{1-3}$$

或

$$F(x) = \int_{-\infty}^{x} f(x)\mathrm{d}x \tag{1-4}$$

所以 X 取值在区间$[\alpha,\beta]$内的概率也可表为

$$P\{\alpha < X < \beta\} = \int_{\alpha}^{\beta} f(x)\mathrm{d}x \tag{1-5}$$

即随机变量落入一区间的概率是以这区间为界的概率密度曲线下的面积(图 1-1)。

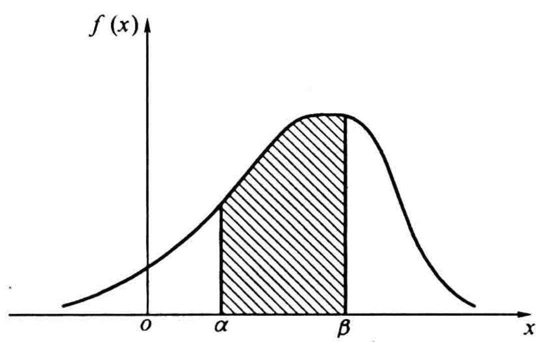

图 1-1

密度曲线 $f(x)$的因次是 $1/x$,其基本特性是:

(1) 它是非负函数,即 $f(x) \geqslant 0$。

(2) $f(x)$ 曲线下的总面积 $\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1$。

分布律完整地描述了随机变量,但在实际工作中常不易获得。而且,在很多场合,常常也只要知道随机变量的基本特性就足够了。这时,可用分布的数字特征。虽然常用的一些数字特征还不足以完整地给出随机变量的变化性质,但它们描述了随机变量的概貌,在实用中有重要意义。

随机变量的**数字特征**得自分布矩。

$$\alpha_k = \mathrm{E}[X^k] = \int_{-\infty}^{\infty} x^k f(x)\mathrm{d}x \tag{1-6}$$

称为 k 阶**原点矩**。$k = 1$ 的一阶原点矩

$$\alpha_1 = \mathrm{E}[X] = \int_{-\infty}^{\infty} x f(x)\mathrm{d}x \tag{1-7}$$

称为随机变量的**均值**或**数学期望**,记以 $\mathrm{E}[X] = m_x$。由于分布曲线 $f(x)$下面积重心的横坐标是 $\int_{-\infty}^{\infty} x f(x)\mathrm{d}x \Big/ \int_{-\infty}^{\infty} f(x)\mathrm{d}x = m_x$,因此均值描述了随机变量分布的位置特征。二阶原点矩

$$\alpha_2 = \int_{-\infty}^{\infty} x^2 f(x)\mathrm{d}x \tag{1-8}$$

称为随机变量的**均方**。均方值的平方根称为均方根(rms)值。

在实际计算中还常应用中心化随机变量,即用随机变量与其均值之差

$$\overset{\circ}{X} = X - m_x \tag{1-9}$$

中心化随机变量的分布矩是中心矩,

$$\mu_k = \mathrm{E}[\overset{\circ}{X}{}^k] = \int_{-\infty}^{\infty} (x - m_x)^k f(x)\mathrm{d}x \tag{1-10}$$

称为 k 阶中心矩。显然,对于任何随机变量,一阶中心矩等于零。二阶中心矩

$$\mu_2 = \mathrm{E}[\overset{\circ}{X}{}^2] = \int_{-\infty}^{\infty} (x - m_x)^2 f(x)\mathrm{d}x \tag{1-11}$$

称为随机变量的**方差**。它在实际计算中具有特殊重要性,记以 $\mathrm{E}[\overset{\circ}{X}{}^2] = \mathrm{D}[X] = D_x$。由(1-11)可见,$D_x$ 为 $f(x)$ 曲线下的面积对于通过 m_x 的轴的惯性矩,因此,它表征了随机变量的可能值相对于均值的离散程度。实用上还常用方差的平方根的绝对值

$$\sigma_x = | D_x^{\frac{1}{2}} | \tag{1-12}$$

它称为随机变量的**标准差**。除了均值和方差这两个基本特征外,有时还用到描述分布不对称程度的偏态系数 $S_x = \mu_3/(D_x)^{3/2}$ 和描述分布峰凸程度(相对于正态分布)的峰态系数 $E_x = \mu_4/D_x^2 - 3$等。

计算中可能要利用原点矩与中心矩之间的关系,较常用的有

$$\mu_2 = \alpha_2 - m_x^2$$

$$\mu_3 = \alpha_3 - 3\alpha_2 m_x + 2m_x^3$$

$$\mu_4 = \alpha_4 - 4\alpha_3 m_x + 6\alpha_2 m_x^3 - 3m_x^4$$

在知道 m_x 和 D_x 而不知 $f(x)$ 时,若需估计随机变量偏离均值超出给定范围 A 的概率,则可利用恰贝舍夫(Chebyshev)不等式得此概率的上界估值

$$P\{|X - m_x| > A\} \leqslant \frac{D_x}{A^2}$$

二、连续随机变量的常用分布

1. 正态分布(高斯 Gauss 分布)

正态分布律是实用中经常遇到的分布律,自然界中很多随机现象都具有这种分布。可以证明,很多个都没有很大影响的独立随机变量之和的分布,趋于正态分布律。

按正态分布的随机变量 X 的概率密度为

$$f(x) = \frac{1}{\sigma_x(2\pi)^{1/2}}\exp\left[-\frac{(x - m_x)^2}{2\sigma_x^2}\right] \tag{1-13}$$

分布函数为

$$F(x) = \frac{1}{2}\left[1 + \Phi\left(\frac{x - m_x}{\sigma_x\sqrt{2}}\right)\right] \tag{1-14}$$

式中,Φ 为拉普拉斯(Laplace)函数,

$$\Phi(u) = \frac{2}{\sqrt{\pi}}\int_0^u \exp(-t^2)\mathrm{d}t$$

可见,正态分布律可由两个参数 m_x(均值)及 σ_x(标准差)完全确定。$f(x)$ 的图象示于

图 1-2。

随机变量 X 出现在区间$[\alpha,\beta]$内的概率为

$$P\{\alpha < X < \beta\} = \frac{1}{2}\left[\Phi\left(\frac{\beta - m_x}{\sigma_x\sqrt{2}}\right) - \Phi\left(\frac{\alpha - m_x}{\sigma_x\sqrt{2}}\right)\right] \tag{1-15}$$

在对称于均值的区间内,由上式可得

$$P\{|X - m_x| < k\sigma_x\} = \Phi\left(\frac{k}{\sqrt{2}}\right)$$

对应于 $k=3$,有 $P=0.997\ 3$。因此,实际计算中常采用 3 倍标准差作为随机变量的极限可能值。

在船舶类问题中,实验证明海浪波面起伏、船舶摇摆角度、波浪中船体的弯曲应力等,都服从正态分布律。正态分布也用于描述材料机械性能(如屈服限)的离散性。

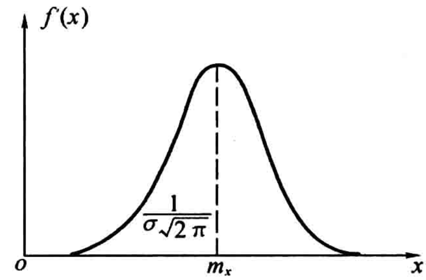

图 1-2

海浪波

的幅值

2. 雷利(Rayleigh)分布

海浪、船舶运动、船体弯曲等可看成是依赖于自变量时间 t 的随机变量 $X(t)$[1]。设 $X(t)$服从正态分布,如 $X(t)$的变化为在每两个零值之间只有一个峰值,且均值 $m_x=0$,如图 1-3(1.2 的四之 2 中将证明这样的 $X(t)$是所谓窄带过程),则 $X(t)$峰(幅)值 R 便服从雷利分布。

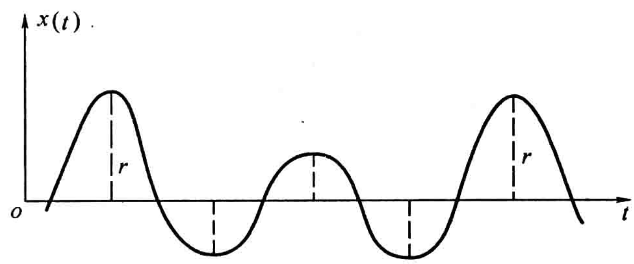

图 1-3

雷利分布的概率密度为

$$f(r) = \frac{2r}{R}\exp\left(-\frac{r^2}{R}\right) \tag{1-16}$$

分布函数为

$$F(r) = 1 - \exp\left(-\frac{r^2}{R}\right) \tag{1-17}$$

可见,雷利分布由单一参数 R 完全确定。$f(r)$的图象如图 1-4。

其均值与标准差分别为

$$m_r = (\pi R)^{\frac{1}{2}}/2 = 0.89\sqrt{R}$$

① 严格说来,$X(t)$是随机过程,图 1-3 是过程的一个现实。但对各态历经过程,图中各峰值便是随机变量 R 的可能值。

$$\sigma_r = \left[\left(1 - \frac{\pi}{4}\right)R\right]^{\frac{1}{2}} = 0.46\sqrt{R} \quad (1-18)$$

常又将 m_r、σ_r 用 $X(t)$ 的方差 D_x 表示，因为 $D_x = \dfrac{R}{2}$（见 1.2 的四之 2），可得

$$\begin{aligned} m_r &= \left(\frac{\pi D_x}{2}\right)^{\frac{1}{2}} = 1.25\sqrt{D_x} \\ \sigma_r^2 &= \left(\frac{4-\pi}{2}\right)D_x = 0.43D_x \end{aligned} \quad (1-19)$$

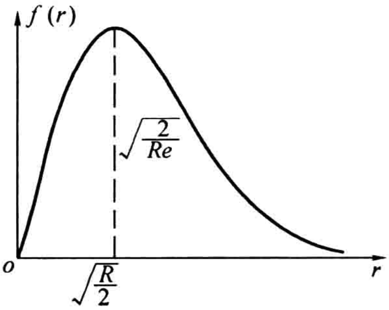

图 1-4

海浪、船舶运动、船体弯曲等，在实用中常可作为窄带过程（短期分布）。因此，它们的幅（峰）值服从雷利分布，这就是雷利分布律在船舶工程中得到广泛应用的原因。

常用于描述不规则海浪的有义波高 $H_{\frac{1}{3}}$ 是指海浪中波高最大的 1/3 部分的平均值。对属于窄带过程的海浪可得 $H_{1/3} = 4\sigma_x$。

3. **赖斯（Rice）分布**

如服从正态分布的 $X(t)$ 的变化为在每两个零值之间可有 n 个峰值图 1-5(a)，即 $X(t)$ 由较小的高频振动叠加在较大的低频振动所组成时，$X(t)$ 的峰值 $\eta = \dfrac{x_{\max}}{\sigma_x}$ 便服从赖斯分布，其概率密度为

$$f(\eta) = \frac{1}{(2\pi)^{\frac{1}{2}}}\left\{\varepsilon\exp\left(-\frac{\eta^2}{2\varepsilon^2}\right) + [2\pi(1-\varepsilon^2)]^{\frac{1}{2}}\eta\exp\left(-\frac{\eta^2}{2}\right)\Phi\left[\frac{(1-\varepsilon^2)^{\frac{1}{2}}}{\varepsilon}\eta\right]\right\} \quad (1-20)$$

其中，带宽系数 ε 可由 $X(t)$ 的图象（现实）统计求得

$$\varepsilon = [1 - (1-2r)^2]^{\frac{1}{2}}$$

而 $r = \dfrac{n^-}{n^+ + n^-}$，其中 n^-、n^+ 分别为负的极大值个数及正的极大值个数。

对应不同 ε 时的 $f(\eta)$ 图象如图 1-5,(b)。注意，$\varepsilon = 0$ 时(1-20)给出了雷利分布，$\varepsilon = 1$时(1-20)给出了均值为零的正态分布（参见 1.2 的四之 2）。

赖斯分布可用于分析宽带过程的海浪、船舶运动、船体弯曲等。

4. **指数分布**

按指数分布的随机变量 X 的概率密度为

$$f(x) = \frac{1}{\lambda}\exp\left(-\frac{x}{\lambda}\right) \quad (1-21)$$

分布函数为

$$F(x) = 1 - \exp\left(-\frac{x}{\lambda}\right) \quad (1-22)$$

所以，指数分布也完全确定于一个参数 λ。$f(x)$ 的图象如图 1-6。

其均值与标准差分别为

$$\begin{aligned} m_x &= \lambda \\ \sigma_x &= \lambda \end{aligned} \quad (1-23)$$

船底砰击压力服从指数分布。实船测量统计表明，指数分布也可用于波浪弯矩峰值的

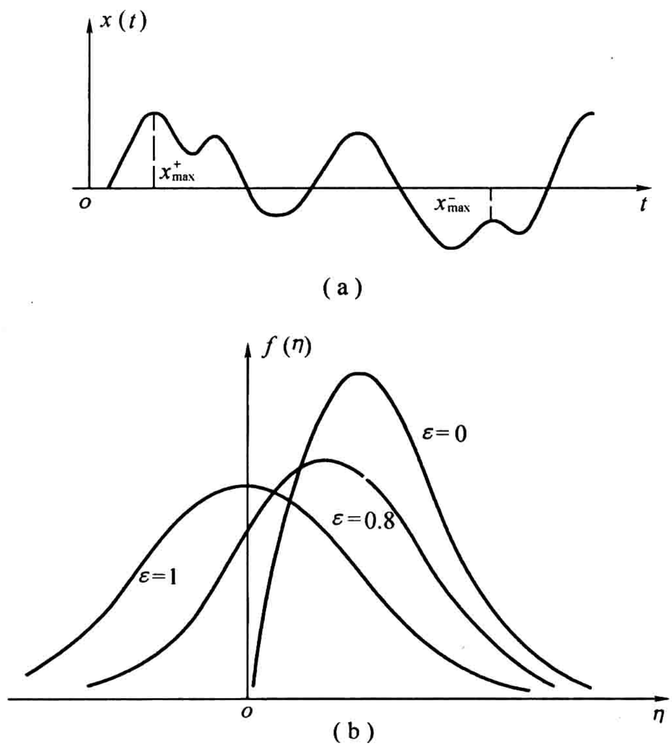

图 1－5

长期分布。

5. 韦布尔(Weibull)分布

这种分布的概率密度为

$$f(x)=\frac{r}{k}\left(\frac{x}{k}\right)^{r-1}\exp\left[-\left(\frac{x}{k}\right)^{r}\right] \quad (1-24)$$

可以看出,这种分布是两参数分布。它的分布函数是

$$F(x)=1-\exp\left[-\left(\frac{x}{k}\right)^{r}\right] \quad (1-25)$$

它的均值与标准差是

$$m_x=k\Gamma\left(1+\frac{1}{r}\right)$$

$$\sigma_x=k\left[\Gamma\left(1+\frac{2}{r}\right)-\Gamma^2\left(1+\frac{1}{r}\right)\right]^{\frac{1}{2}} \quad (1-26)$$

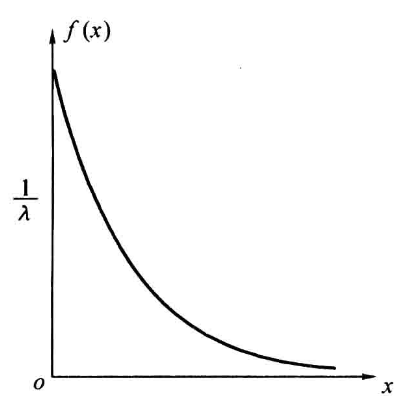

图 1－6

其中,伽马(Γamma)函数为

$$\Gamma(u)=\int_0^{\infty}t^{u-1}\exp(-t^2)\mathrm{d}t$$

若韦布尔分布中的一个参数取值 $r=2$,则(1－24)成为

$$f(x)=\frac{2x}{k^2}\exp\left(-\frac{x^2}{k^2}\right)$$

这便是雷利分布,其中 $k=\sqrt{R}$。若 $r=1$,则(1－24)成为

$$f(x) = \frac{1}{k}\exp\left(-\frac{x}{k}\right)$$

这便是指数分布,其中 $k=\lambda$。因此,韦布尔分布可用于分析波浪弯矩与船体应力等的短期分布以及长期分布①。

6. **对数正态分布**

有些实测表明,海浪高度(一年以上观测数据)、船舶摇摆角度和船体构件应力的长期分布服从对数正态分布。设 X 为具有参数(m,s)的对数正态分布,则其概率密度为

$$f(x) = \frac{1}{(2\pi)^{\frac{1}{2}}sx}\exp\left[-\frac{1}{2}\left(\frac{\lg x - m}{s}\right)^2\right] \tag{1-27}$$

其均值与方差是

$$\begin{aligned} m_x &= \exp\left(m+\frac{s^2}{2}\right) \\ D_x &= \exp(2m+2s^2) - \exp(2m+s^2) \end{aligned} \tag{1-28}$$

7. **反余弦分布**

若随机变量为

$$X = a\cos\theta$$

其中 a ——非随机幅值;

θ ——在区间 $0\leqslant\theta\leqslant\pi$ 内均匀分布的随机(自)变量。

则 X 的概率密度及分布函数分别为

$$f(x) = \begin{cases} 0 & \text{当 } x<-a \\ \dfrac{1}{\pi(a^2-x^2)^{\frac{1}{2}}} & \text{当 } -a<x<a \\ 0 & \text{当 } x>a \end{cases} \tag{1-29}$$

$$F(x) = \begin{cases} 0 & \text{当 } x<-a \\ 1-\dfrac{1}{\pi}\cos^{-1}\dfrac{x}{a} & \text{当 } -a<x<a \\ 0 & \text{当 } x>a \end{cases} \tag{1-30}$$

它的均值与标准差是

$$\begin{aligned} m_x &= 0 \\ \sigma_x &= \frac{a}{\sqrt{2}} \end{aligned} \tag{1-31}$$

这种分布出现于用三角级数模型模拟正态过程时(参见 1.3 的五之 3)。

8. **χ^2 分布**(皮尔逊 Pierson 分布)

按 χ^2 分布的随机变量 X 的概率密度为

$$f(x) = \frac{1}{2^{r/2}\Gamma\left(\frac{r}{2}\right)}x^{\frac{r}{2}-1}\exp\left(-\frac{x}{2}\right) \tag{1-32}$$

分布函数为

① 短期是指统计时间大约为半小时至数小时,长期是指统计时间为一年或更长。

$$F(x) = \frac{1}{2^{\frac{r}{2}}\Gamma\left(\frac{r}{2}\right)}\int_0^x x^{\frac{r}{2}-1}\exp\left(-\frac{x}{2}\right)\mathrm{d}x \tag{1-33}$$

其中,参数 r 称为自由度数。它的均值与标准差是

$$\begin{aligned} m_x &= r \\ \sigma_x &= (2r)^{\frac{1}{2}} \end{aligned} \tag{1-34}$$

皮尔逊分布函数 $P\{\chi^2, r\} = P\{X > \chi^2\} = 1 - F(\chi^2)$常用于检验实验分布律与理论分布律的拟合程度(参见本节的五)。

三、随机变量系

在实际问题中,实验结果常须由相当多的随机变量来描述(例如船舶的运动同时确定于海浪浪级、浪向、船舶装载、航速、航向角等),这些随机变量就构成了随机变量系。现在讨论可能值为 x、y 的两个变量 X、Y 构成的最简单的随机变量系。

为完整地描述随机变量系的概率特性,需要知道它的分布律。随机变量系(X, Y)的分布函数是

$$F(x,y) = P\{X < x, Y < y\} \tag{1-35}$$

其密度函数是

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y} \tag{1-36}$$

分布函数用密度函数表达时为

$$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(x,y)\mathrm{d}x\mathrm{d}y \tag{1-37}$$

X 出现在区间$[\alpha, \beta]$内,同时 Y 出现在区间$[\gamma, \delta]$内的概率为

$$P\{\alpha < X < \beta, \gamma < Y < \delta\} = \int_{\alpha}^{\beta}\int_{\gamma}^{\delta} f(x,y)\mathrm{d}x\mathrm{d}y \tag{1-38}$$

知道系的分布,可以容易地得到系中各变量的分布(称为边缘分布)

$$\begin{aligned} f_1(x) &= \int_{-\infty}^{\infty} f(x,y)\mathrm{d}y \\ f_2(y) &= \int_{-\infty}^{\infty} f(x,y)\mathrm{d}x \end{aligned} \tag{1-39}$$

但是,要从系中各变量的分布求系的分布时,还必须知道系中各变量之间的概率关系。这种关系用所谓**条件分布**表达。若记 $f_1(x \mid y)$和 $f_2(y \mid x)$分别为 X 在条件 $Y = y$ 下和 Y 在条件 $X = x$ 下的条件概率密度,则系的密度函数可表示为

$$\begin{aligned} f(x,y) &= f_1(x)f_2(y \mid x) \\ f(x,y) &= f_2(y)f_1(x \mid y) \end{aligned} \tag{1-40}$$

如果随机变量 X 与 Y 相互独立(即其中每一个的分布律不依赖于另一个的取值),则有

$$f(x,y) = f_1(x)f_2(y) \tag{1-41}$$

即系的分布密度等于系中各随机变量的分布密度的乘积。

在实际问题中,要得到随机变量系的分布律常是很困难的,同时也常是不必要的。因此,常用描述系的基本特性的数字特征。这些数字特征也得自分布矩。

$$\alpha_{ks} = \mathrm{E}[X^k Y^s] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^k y^s f(x,y)\mathrm{d}x\mathrm{d}y \tag{1-42}$$

称为变量系(X,Y)的 k、s 阶原点矩。$\alpha_{10}=m_x,\alpha_{01}=m_y$ 表示各随机变量的均值。

$$\mu_{ks}=\mathrm{E}[\mathring{X}^k\mathring{Y}^s]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-m_x)^k(y-m_y)^s f(x,y)\mathrm{d}x\mathrm{d}y \tag{1-43}$$

称为(X,Y)的 k、s 阶中心矩。$\mu_{20}=D_x,\mu_{02}=D_y$ 表示各随机变量的方差。

均值和方差没有表明变量之间的关系,因此,还必须用表征变量间相关性的数字特征,这便是二阶混合矩。二阶混合原点矩

$$\alpha_{11}=\mathrm{E}[XY]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyf(x,y)\mathrm{d}x\mathrm{d}y \tag{1-44}$$

称为**相关矩**。二阶混合中心矩

$$\mu_{11}=\mathrm{E}[\mathring{X}\mathring{Y}]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-m_x)(y-m_y)f(x,y)\mathrm{d}x\mathrm{d}y \tag{1-45}$$

称为**协方差**。显然,均值为零时,相关矩与协方差相同,记 $\mu_{11}=K_{xy}$。为比较变量间的相关程度,通常用无因次的**相关系数**

$$r_{xy}=\frac{K_{xy}}{\sigma_x\sigma_y}\qquad -1\leqslant r_{xy}\leqslant 1 \tag{1-46}$$

若 X 与 Y 相互独立,显然 $r_{xy}=0$。

上述全部内容都可以推广应用于任意多个随机变量组成的系。

四、随机变量函数

在实践中所遇到的物理量常是有函数关系的,例如,作用在船体上的水压力可认为是与船舶对于海浪的相对速度平方成正比的。因此,常需要由已知自变量的分布律推算其函数的分布律。这样的间接方法便构成了概率分析的主要内容(参见 1.3)。

设已知随机变量 X 的密度函数为$f(x)$,而 Y 是X 的给定函数 $Y=\varphi(X)$,需求 Y 的概率密度 $g(y)$。

设 φ 为非单调函数,将 X 的变化区间划成很多小段,使每一段内函数 φ_i 是单调的,如

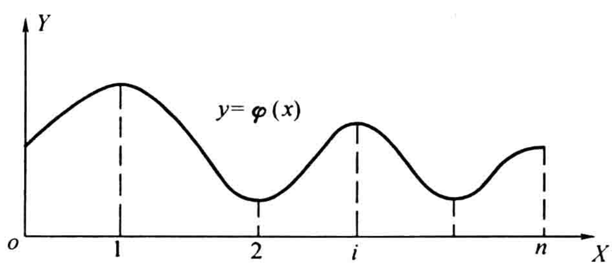

图 1-7

图 1-7。记每一小段内 φ_i 的逆函数为 $X=\psi_i(Y)$,则可得

$$g(y)=\sum_{i=1}^{n}f[\psi_i(y)]\,|\psi_i'(y)| \tag{1-47}$$

例如给定 $Y=X^2$,则由上式可得

$$g(y) = \frac{f(\sqrt{y}) + f(-\sqrt{-y})}{2\sqrt{y}}$$

当 X 为雷利分布时，得 Y 为指数分布；当 X 为正态分布时，得 Y 为单自由度 χ^2 分布。一般来说，随机变量函数的分布不同于随机变量本身的分布。

正态分布随机变量的一个重要特征是它的线性变换仍是正态变量。设

$$f(x) = \frac{1}{\sigma_x (2\pi)^{\frac{1}{2}}} \exp\left[-\frac{(x - m_x)^2}{2\sigma_x^2}\right]$$

且

$$y = ax + b$$

则有

$$g(y) = \frac{1}{|a|\sigma_x (2\pi)^{\frac{1}{2}}} \exp\left[-\frac{[y - (am_x + b)]^2}{2|a|^2\sigma_x^2}\right]$$

这仍是正态分布，其参数为

$$m_y = am_x + b$$
$$\sigma_y = |a|\sigma_x$$

正态分布的这一特性，使它在造船问题中得到广泛应用。即如船舶可认为是线性系统，则其输入（海浪）为正态变量时，其输出（运动、应力）也是正态分布的。

若需要计算的是随机变量函数 Y 的数字特征，则不必知道函数的概率分布，只需要知道随机变量 X 本身的概率分布就够了。当 $Y = \varphi(X)$时，有

$$\begin{aligned} m_y &= \int_{-\infty}^{\infty} \varphi(x) f(x) \mathrm{d}x \\ D_y &= \int_{-\infty}^{\infty} [\varphi(x)]^2 f(x) \mathrm{d}x - m_y^2 \end{aligned} \tag{1-48}$$

由上式可得有关数字特征的下列定理：

(1) 随机变量与非随机变量的乘积

若
$$Y = aX$$
则
$$m_y = am_x$$
$$D_y = a^2 D_x$$

(2) 随机变量的和

若
$$Y = X_1 + X_2$$
则
$$\begin{aligned} m_y &= m_{x_1} + m_{x_2} \\ D_y &= D_{x_1} + D_{x_2} + 2K_{x_1 x_2} \\ &= D_{x_1} + D_{x_2} + 2r_{x_1 x_2}(D_{x1} D_{x2})^{\frac{1}{2}} \end{aligned}$$

(3) 随机变量的线性函数

若
$$Y = \sum_i a_i X_i + b$$
则
$$m_y = \sum_i a_i m_{x_i} + b$$

$$D_y = \sum_i a_i^2 D_{x_i} + 2\sum_{i<j} a_i a_j K_{ij}$$

(4) 随机变量的乘积

若
$$Y = X_1 X_2$$
则
$$m_y = m_{x_1} m_{x_2}$$
$$D_y = D_{x_1} m_{x_2}^2 + D_{x_2} m_{x_1}^2 + D_{x_1} D_{x_2}$$

对于随机变量系(X_1, X_2),若已知其概率密度$f(x_1, x_2)$,需求函数关系为

$$Y_1 = \varphi_1(X_1, X_2)$$
$$Y_2 = \varphi_2(X_1, X_2)$$

的函数(Y_1, Y_2)的概率分布时,用逆变换

$$X_1 = \psi_1(Y_1, Y_2)$$
$$X_2 = \psi_2(Y_1, Y_2)$$

可得

$$g(y_1, y_2) = f[\psi_1(y_1, y_2), \psi_2(y_1, y_2)] I$$

式中,雅可比变换

$$I = \begin{vmatrix} \dfrac{\partial \psi_1}{\partial y_1} & \dfrac{\partial \psi_1}{\partial y_2} \\ \dfrac{2\psi_2}{\partial y_1} & \dfrac{\partial \psi_2}{\partial y_2} \end{vmatrix} = \begin{vmatrix} \dfrac{\partial \varphi_1}{\partial x_1} & \dfrac{\partial \varphi_1}{\partial x_2} \\ \dfrac{\partial \varphi_2}{\partial x_1} & \dfrac{\partial \varphi_2}{\partial x_2} \end{vmatrix}^{-1}$$

五、随机变量实验分布律

设随机变量 X 的实验结果按其数值大小排列为$x_1, x_2, \cdots, x_n$。将实验值的总范围分成若干区间(通常分 10~20 区间),并计算 X 出现于各区间的频率

$$p_i^* = \frac{m_i}{n} \tag{1-49}$$

其中,m_i 是实验值出现于第 i 区间的个数,n 是实验值总数。用这个频率与实验值范围的关系可作表示变量分布概貌的**直方图**,如图 1-8,图中每一矩形面积等于 X 出现于该区间的频率p_i^*。

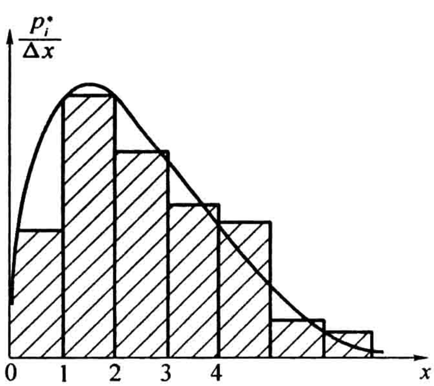

图 1-8

为对变量 X 进行概率分析,需对此实验分布拟合一理论分布曲线。理论分布曲线的类型应当根据变量的实际物理本质选定。理论分布曲线解析式中所包含的参数,须选择得使该曲线能最优地描述直方图表述的实验分布。这个最优配合问题常采用矩的方法解决,即取理论分布的数字特征等于相应的实验统计特征。

求实验分布数字特征时,如果实验结果个数 n 足够大,则可认为每一区间内的实验值 x_i 是等于该区间平均值的常数。故由实验结果可得均值和方差为