教学内容

§3-3 条件分布

本节重点

一、离散型随机变量的条件分布

二、二维随机变量条件分布函数

三、连续型随机变量条件概率密度

四、例题分析

五、思考与小结

知识点回顾

在第一章中，我们介绍了条件概率的概念.

在事件B发生的条件下事件A发生的条件概率

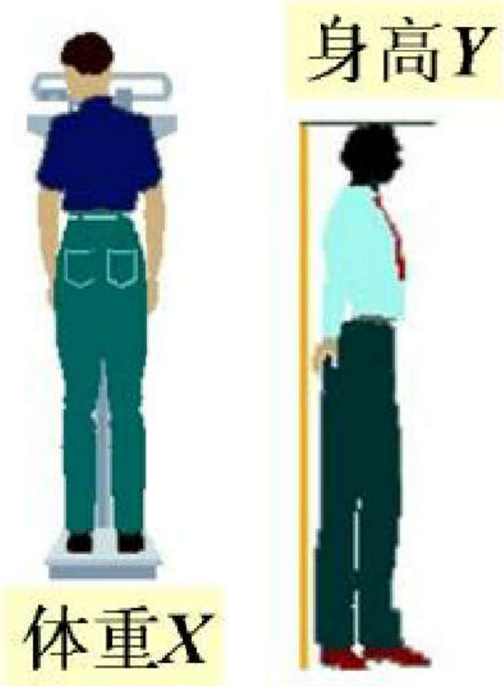
$$P(A|B)=\frac{P(AB)}{P(B)}$$

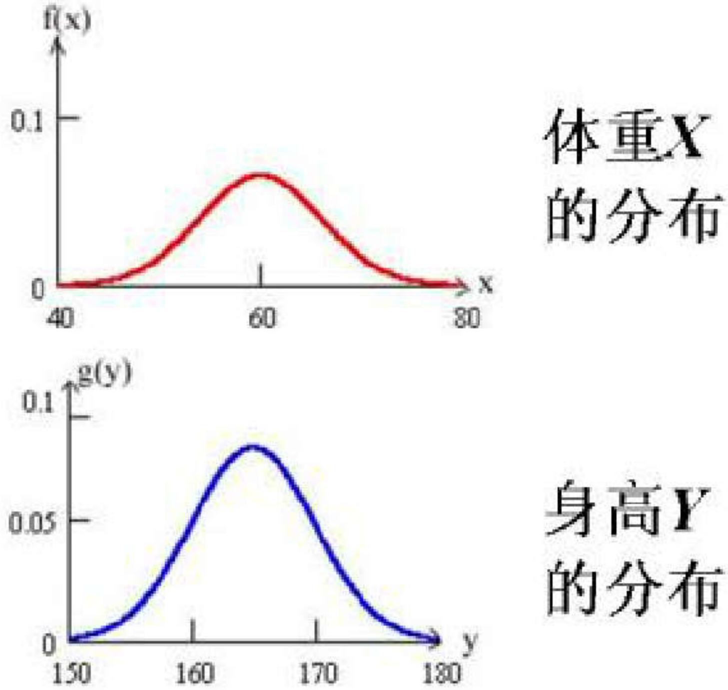
推广到随机变量

设有两个$r.v\ X,Y$， 在给定Y取某个或某些值的条件下，求X的概率分布.

这个分布就是条件分布.

例如，考虑某大学的全体学生，从其中随机抽取一个学生，分别以X和Y表示其体重和身高，则X和Y都是随机变量，它们都有一定的概率分布.





现在若限制$1.7<Y<1.8$（米），在这个条件下去求X的条件分布，这就意味着要从该校的学生中把身高在1.7米和1.8米之间的那些人都挑出来，然后在挑出的学生中求其体重的分布.

容易想象，这个分布与不加这个条件时的分布会很不一样.

例如，在条件分布中体重取大值的概率会显著增加 .

一、离散型r.v的条

类似定义在$X=x_i$条件下
随机变量Y的条件概率函数.

实际上是第一章

一种形式下的体现.

定义1 设(X,Y)是二维离散型随机变量，对于固定的j，若$P(Y=y_j)>0$，则称——**67页**

$$P(X=x_i|Y=y_j)=\frac{P(X=x_i,Y=y_j)}{P(Y=y_j)}=\frac{p_{ij}}{p_{\bullet j}},\ i=1,2,\ldots$$

称为在$Y=y_j$条件下随机变量X的条件概率函数.

作为条件的那个r.v,认为取值是给定的，在此条件下求另一r.v的概率分布.

条件分布是一种概率分布，它具有概率分布的一切性质．正如条件概率是一种概率，具有概率的一切性质．

例如：

$$P(X=x_i \mid Y=y_j) \geq 0,$$

i=1,2, …

$$\sum_{i=1}^{\infty} P(X=x_i \mid Y=y_j)=1$$

例1 在一汽车工厂中，一辆汽车有两道工序是由机器人完成的. 其一是紧固3只螺栓，其二是焊接2处焊点. 以X表示螺栓紧固得不良的数目，以Y表示由机器人焊接的不良焊点的数目. 据积累的资料知(X,Y)具有分布律：

$Y \backslash X$	0	1	2	3	$P\{Y=j\}$
0	0.840	0.030	0.020	0.010	0.900
1	0.060	0.010	0.008	0.002	0.080
2	0.010	0.005	0.004	0.001	0.020
$P\{X=i\}$	0.910	0.045	0.032	0.013	1.000

(1) 求在 $X=1$ 的条件下, Y 的条件分布律;

(2) 求在 $Y=0$ 的条件下, X 的条件分布律.

解 边缘分布已经求出-----列在上表中.

在 $X=1$ 的条件下, Y 的条件分布律为

$$P\{Y=0|X=1\}=\frac{P\{X=1,Y=0\}}{P\{X=1\}}=\frac{0.030}{0.045},$$

$$P\{Y=1|X=1\}=\frac{P\{X=1,Y=1\}}{P\{X=1\}}=\frac{0.010}{0.045},$$

$$P\{Y=2|X=1\}=\frac{P\{X=1,Y=2\}}{P\{X=1\}}=\frac{0.005}{0.045},$$

或写成

$Y=k$	0	1	2
$P\{Y=k\|X=1\}$	$\frac{6}{9}$	$\frac{2}{9}$	$\frac{1}{9}$

同样可得在$Y=0$的条件下X的条件分布律为

$X=k$	0	1	2	3
$P\{X=k\|Y=0\}$	$\frac{84}{90}$	$\frac{3}{90}$	$\frac{2}{90}$	$\frac{1}{90}$

二、连续型$r.v$的条件分布

设(X,Y)是二维连续型$r.v$，由于对任意x, y, $P(X=x)=0, P(Y=y)=0$，所以不能直接用条件概率公式得到条件分布，下面我们直接给出条件概率密度的定义.

•教材69页—3.4式

定义 设X和Y的联合概率密度为$f(x,y)$，边缘概率密度为$f_X(x), f_Y(y)$，则对一切使$f_X(x)>0$的x，定义已知$X=x$条件下，Y的条件密度函数为

$$f_{Y|X}(y\,|\,x)=\frac{f(x,y)}{f_X(x)}$$

同样，对一切使$f_Y(y)>0$的y，定义

$$f_{X|Y}(x\,|\,y)=\frac{f(x,y)}{f_Y(y)}$$

称为已知$Y=y$条件下，X的条件密度函数．

•教材69页—3.4式

运用条件概率密度，我们可以在已知某一随机变量值的条件下，定义与另一随机变量有关的事件的条件概率.

即：若(X,Y)是连续型$r.v$，则对任一集合A，

$$P(X \in A \mid Y = y) = \int_A f_{X|Y}(x \mid y)dx$$

特别，取 $A = (-\infty, u)$，

定义在已知 $Y=y$下，X的条件分布函数为

$$F_{X|Y}(x \mid y) = P(X \le u \mid Y = y)$$

$$= \int_{-\infty}^{u} f_{X|Y}(x \mid y)dx$$

三、条件分布函数与条件密度函数的关系

$$F_{X|Y}(x|y) = P(X \le x|Y = y)$$

$$= \int_{-\infty}^{x} f_{X|Y}(x|y)\,\mathrm{d}\,x$$

$$= \int_{-\infty}^{x} [f(x,y)/f_Y(y)]\,\mathrm{d}\,x.$$

$$F_{Y|X}(y|x) = \int_{-\infty}^{y} f_{Y|X}(y|x)\,\mathrm{d}\,y$$

$$= \int_{-\infty}^{y} [f(x,y)/f_X(x)]\,\mathrm{d}\,y.$$

•具体计算条件概率的方法

$$P(X \le x|Y < y) = ?$$

说明

联合分布、边缘分布、条件分布的关系如下:

$$P(X \le x \mid Y < y) = \frac{P(X \le x, Y < y)}{P(Y < y)}$$

$$= \frac{\iint_G p(x, y)\,dx\,dy}{\int_L p(y)\,dy}$$

例2 设(X,Y)服从单位圆上的均匀分布，概率密度为

$$f(x,y)=\begin{cases}\dfrac{1}{\pi}, & x^2+y^2\le 1\\ 0, & \text{其它}\end{cases}$$

求$f_{Y|X}(y|x)$

解 X的边缘密度为

$$f_X(x)=\int_{-\infty}^{\infty}f(x,y)dy=\begin{cases}\dfrac{2}{\pi}\sqrt{1-x^2}, & |x|\le 1\\ 0, & |x|>1\end{cases}$$

当$|x|<1$时,有

$$f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)}=\frac{1/\pi}{(2/\pi)\sqrt{1-x^2}}$$

$$=\frac{1}{2\sqrt{1-x^2}},\quad -\sqrt{1-x^2}\le y\le\sqrt{1-x^2}$$

即当$|x|<1$时，有

X作为已知变量

$$f_{Y|X}(y\mid x)=\begin{cases}\dfrac{1}{2\sqrt{1-x^2}}, & -\sqrt{1-x^2}\le y\le\sqrt{1-x^2}\\ 0, & \text{其它}\end{cases}$$

X已知下Y的条件密度

这里是y的取值范围

考研题 设数X在区间(0,1)均匀分布，当观察到$X=x(0<x<1)$时，数Y在区间$(x,1)$上随机地取值. 求Y的边缘概率密度.

解 依题意，X具有概率密度

$$f_X(x)=\begin{cases}1, & 0<x<1\\ 0, & 其它\end{cases}$$

对于任意给定的值$x(0<x<1)$，在$X=x$的条件下，Y的条件概率密度为

$$f_{Y|X}(y|x)=\begin{cases}\dfrac{1}{1-x}, & x<y<1\\ 0, & 其它\end{cases}$$

X和Y的联合密度为

$$f(x,y)=f_X(x)f_{Y|X}(y|x)$$

已知边缘密度、条件密度，求联合密度

$$=\begin{cases}\dfrac{1}{1-x}, & 0<x<y<1\\ 0, & \text{其它}\end{cases}$$

于是得Y的边缘概率密度为

$$f_Y(y)=\int_{-\infty}^{\infty}f(x,y)dx$$

$$=\begin{cases}\displaystyle\int_0^y\frac{1}{1-x}dx=-\ln(1-y), & 0<y<1\\ 0, & \text{其它}\end{cases}$$

内容小结

1. 设 (X,Y) 是二维离散型随机变量 $p_{ij}(i,j=1,2\cdots)$ 为其联合分布律，在给定 $Y=y_j$ 条件下随机变量 X 的条件分布律为

$$P\{X=x_i|Y=y_j\}=\frac{P\{X=x_i,Y=y_j\}}{P\{Y=y_j\}}=\frac{p_{ij}}{p_{\bullet j}},$$

在给定 $X=x_i$ 条件下随机变量 Y 的条件分布律为

$$P\{Y=y_j|X=x_i\}=\frac{P\{X=x_i,Y=y_j\}}{P\{X=x_i\}}=\frac{p_{ij}}{p_{i\bullet}},$$

其中 $i,j=1,2,\cdots$.

2. 设 (X,Y) 是二维连续型随机变量则有

$$F_{X|Y}(x|y) = \int_{-\infty}^{x} f_{X|Y}(x|y)\,\mathrm{d}x$$

$$= \int_{-\infty}^{x} [f(x,y)/f_Y(y)]\,\mathrm{d}x.$$

$$F_{Y|X}(y|x) = \int_{-\infty}^{y} f_{Y|X}(y|x)\,\mathrm{d}y$$

$$= \int_{-\infty}^{y} [f(x,y)/f_X(x)]\,\mathrm{d}y.$$

考研题　设随机变量,(X,Y)的概率密度为

$$f(x,y)=\begin{cases} e^{-y}, & 0<x<y \\ 0, & \text{其它} \end{cases}$$

(1)求 X与Y 的边际概率密度，并判断是否 独立

(2)求在$Y=y$的条件下，X的条件概率密度

(3)求概率$P(X+2Y\le 1)$;

$P(X\ge 2\mid Y<4)$; $P(X\ge 2\mid Y=4)$

分析: $f_X(x)=\int_{-\infty}^{\infty} f(x,y)dy$

$$f_X(x)=\begin{cases} e^{-x}, & x>0 \\ 0, & x\le 0 \end{cases} \quad f_Y(y)=\begin{cases} ye^{-y}, & y>0 \\ 0, & y\le 0 \end{cases}$$

$$f(x,y)\ne f_X(x)f_Y(y)$$

考研题　设随机变量,(X,Y)的概率密度为

(1)求 X与Y 的边际概率密度，并判断是否 独立

$$f(x,y)=\begin{cases} e^{-y}, & 0<x<y \\ 0, & 其它 \end{cases}$$

(2)求在$Y=y$的条件下，X的条件概率密度

(3)求概率$P(X+2Y\le 1)$;

$P(X\ge 2\mid Y<4)$; $P(X\ge 2\mid Y=4)$

分析:

当$y>0$时　$$f_{X|Y}(x\mid y)=\frac{f(x,y)}{f_Y(y)}=\frac{e^{-y}}{ye^{-y}}=\frac{1}{y}$$

$$f_{X|Y}(x\mid y)=\begin{cases} \dfrac{1}{y} & 0<x<y \\ 0 & 其它 \end{cases}$$

考研题　设随机变量,(X,Y)的概率密度为

$$f(x,y)=\begin{cases} e^{-y}, & 0<x<y \\ 0\ , & \text{其它} \end{cases}$$

(1)求 X与Y 的边际概率密度，并判断是否 独立

(2)求在$Y=y$的条件下，X的条件概率密度

(3)求概率$P(X+2Y\le 1)$;

$P(X\ge 2\mid Y<4)$; $P(X\ge 2\mid Y=4)$

分析:求概率$P(X+2Y\le 1)$

$$=\iint\limits_{X+2Y\le 1} f(x,y)dxdy$$

$$=\int_0^{\frac{1}{3}}dx\int_x^{\frac{1-x}{2}}e^{-y}dy=1+2e^{-\frac{1}{2}}-3e^{-\frac{1}{3}}$$

考研题　设随机变量,(X,Y)的概率密度为

$$f(x,y)=\begin{cases} e^{-y}, & 0<x<y \\ 0, & 其它 \end{cases}$$

(1)求 X与Y 的边际概率密度，并判断是否独立

(2)求在$Y=y$的条件下，X的条件概率密度

(3)求概率$P(X+2Y\le 1)$;

$P(X\ge 2|Y<4)$; $P(X\ge 2|Y=4)$

分析:求概率$P(X\ge 2|Y<4)$

$$=\frac{P(X>2,Y<4)}{P(Y<4)}=\frac{\int_2^4 dy\int_2^y e^{-y}dx}{\int_0^4 ye^{-y}dy}$$

考研题 设随机变量,(X,Y) 的概率密度为

$$f(x,y)=\begin{cases} e^{-y}, & 0<x<y \\ 0\ , & \text{其它} \end{cases}$$

(1)求 X与Y 的边际概率密度，并判断是否 独立

(2)求在$Y=y$ 的条件下， X的条件概率密度

(3)求概率$P\ (X+2Y\le 1)$;

$P\ (X\ge 2\,|\,Y<4)$; $P\ (X\ge 2\,|\,Y=4)$

分析:求概率$P\ (X\ge 2\,|\,Y=4)$ $\quad P(Y=4)=0$

$$f_{X|Y}(x\,|\,y)=\begin{cases} \dfrac{1}{y} & 0<x<y \\ 0 & \text{其它} \end{cases} \Rightarrow f_{X|Y}(x\,|\,4)=\begin{cases} \dfrac{1}{4} & 0<x<4 \\ 0 & \text{其它} \end{cases}$$

$$P\ (X\ge 2\,|\,Y=4)=\int_2^4\frac{1}{4}dx=\frac{1}{2}$$

例 设(X,Y)的概率密度是

$$f(x,y)=\begin{cases}\dfrac{e^{-x/y}e^{-y}}{y}, & 0<x<\infty, 0<y<\infty \\ 0, & \text{其它}\end{cases}$$

求 $P(X>1|Y=y)$

解: $P(X>1|Y=y)=\int_1^{\infty} f_{X|Y}(x|y)dx$

为此, 需求出 $f_{X|Y}(x|y)$

由于 $f_Y(y)=\int_{-\infty}^{\infty} f(x,y)dx$

$$=\int_0^{\infty}\frac{e^{-x/y}e^{-y}}{y}dx=\frac{e^{-y}}{y}\left[-ye^{-x/y}\right]\Big|_0^{\infty}=e^{-y},$$

$$0<y<\infty$$

于是对$y>0$,

$$f_{X|Y}(x\mid y)=\frac{f(x,y)}{f_Y(y)}=\frac{e^{-x/y}}{y},\quad x>0$$

故对$y>0$, $P(X>1|Y=y)=\int_1^{\infty}\frac{e^{-x/y}}{y}dx$

$$=-e^{-x/y}\Big|_1^{\infty}=e^{-1/y}$$