3.1.6 Conditional expectations

3.1.6 Conditional expectations

Conditional distribution:

3.1.6 Conditional expectations

Conditional distribution:

In the discrete case, the conditional probability mass function of ξ for given η is

$$p_{\xi|\eta}(x_i|y_j) = \frac{P(\xi = x_i, \eta = y_j)}{P(\eta = y_j)} = \frac{p_{ij}}{p_{\cdot j}}.$$

3.1.6 Conditional expectations

Conditional distribution:

In the discrete case, the conditional probability mass function of ξ for given η is

$$p_{\xi|\eta}(x_i|y_j) = \frac{P(\xi = x_i, \eta = y_j)}{P(\eta = y_j)} = \frac{p_{ij}}{p_{\cdot j}}.$$

In the continuous case, the conditional probability density function of ξ for given η is

$$p_{\xi|\eta}(x|y) = \frac{p(x,y)}{p_\eta(y)}.$$

In general, if the limit

$$F_{\xi|\eta}(x|y) = \lim_{\epsilon\to 0} \frac{P(\xi \le x, -\epsilon + y < \eta < \epsilon + y)}{P(-\epsilon + y < \eta < \epsilon + y)}$$

exists for all x, we call it the conditional distribution of ξ for given $\eta = y$.

The conditional distribution possesses the same properties as usual distribution functions. Therefore we can calculate mathematical expectations with respect to it.

The conditional distribution possesses the same properties as usual distribution functions. Therefore we can calculate mathematical expectations with respect to it.

The mathematical expectation of a conditional distribution is called the conditional mathematical expectation:

$$E[\eta|\xi = x] = \int_{-\infty}^{\infty} y dF_{\eta|\xi}(y|x).$$

为了强调y是积分变量,上述积分也常写为$\int_{-\infty}^{\infty} yF_{\eta|\xi}(dy|x)$.

If η has conditional density $p_{\eta|\xi}(y|x)$ given $\xi = x$, then

$$E(\eta|\xi = x) = \int_{-\infty}^{+\infty} y p_{\eta|\xi}(y|x) dy.$$

If η has conditional density $p_{\eta|\xi}(y|x)$ given $\xi = x$, then

$$E(\eta|\xi = x) = \int_{-\infty}^{+\infty} y p_{\eta|\xi}(y|x) dy.$$

Example

$(\xi, \eta) \sim N(a, b, \sigma_1^2, \sigma_2^2, r)$, then

$\eta\big|_{\xi=x} \sim$

If η has conditional density $p_{\eta|\xi}(y|x)$ given $\xi = x$, then

$$E(\eta|\xi = x) = \int_{-\infty}^{+\infty} y p_{\eta|\xi}(y|x) dy.$$

Example

$(\xi, \eta) \sim N(a, b, \sigma_1^2, \sigma_2^2, r)$, then

$\eta|_{\xi=x} \sim N(b + r\sigma_2(x - a)/\sigma_1, \sigma_2^2(1 - r^2))$.

If η has conditional density $p_{\eta|\xi}(y|x)$ given $\xi = x$, then

$$E(\eta|\xi = x) = \int_{-\infty}^{+\infty} y p_{\eta|\xi}(y|x)dy.$$

Example

$(\xi, \eta) \sim N(a, b, \sigma_1^2, \sigma_2^2, r)$, then

$\eta|_{\xi=x} \sim N(b + r\sigma_2(x-a)/\sigma_1, \sigma_2^2(1-r^2))$.

In turn,

$$E(\eta|\xi = x) = b + r\frac{\sigma_2}{\sigma_1}(x-a).$$

Denote by $E(\eta|\xi)$: when $\xi = x$ the function takes value $E(\eta|\xi = x)$. $E(\eta|\xi)$ is a r.v. and a function of ξ.

$$E(E(\eta|\xi)) = E\eta.$$

Proof.

Denote by $E(\eta|\xi)$: when $\xi = x$ the function takes value $E(\eta|\xi = x)$. $E(\eta|\xi)$ is a r.v. and a function of ξ.

$$E(E(\eta|\xi)) = E\eta.$$

Proof. We give the proof only for continuous random variables below. Suppose that (ξ, η) has the pdf $p(x, y)$. In this case,

$$p_\xi(x) = \int_{-\infty}^{\infty} p(x, y)dy,$$

and then

and then

$$p_{\eta|\xi}(y|x) = \frac{p(x,y)}{p_\xi(x)}, \quad (p_\xi(x) \neq 0),$$

and then

$$p_{\eta|\xi}(y|x) = \frac{p(x,y)}{p_\xi(x)}, \quad (p_\xi(x) \neq 0),$$

$$E(\eta|\xi = x) = \int_{-\infty}^{\infty} y \frac{p(x,y)}{p_\xi(x)} dy.$$

and then

$$p_{\eta|\xi}(y|x) = \frac{p(x,y)}{p_\xi(x)}, \quad (p_\xi(x) \neq 0),$$
$$E(\eta|\xi = x) = \int_{-\infty}^{\infty} y \frac{p(x,y)}{p_\xi(x)} dy.$$

So

$$E(E(\eta|\xi)) \;=\; \int_{-\infty}^{\infty} E(\eta|\xi = x) p_\xi(x) dx$$

and then

$$p_{\eta|\xi}(y|x) = \frac{p(x,y)}{p_\xi(x)}, \quad (p_\xi(x) \neq 0),$$

$$E(\eta|\xi = x) = \int_{-\infty}^{\infty} y \frac{p(x,y)}{p_\xi(x)} dy.$$

So

$$\begin{aligned} E(E(\eta|\xi)) &= \int_{-\infty}^{\infty} E(\eta|\xi = x) p_\xi(x) dx \\ &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} y \frac{p(x,y)}{p_\xi(x)} dy \right) p_\xi(x) dx \end{aligned}$$

and then

$$p_{\eta|\xi}(y|x) = \frac{p(x,y)}{p_\xi(x)}, \quad (p_\xi(x) \neq 0),$$

$$E(\eta|\xi = x) = \int_{-\infty}^{\infty} y \frac{p(x,y)}{p_\xi(x)} dy.$$

So

$$\begin{aligned} E(E(\eta|\xi)) &= \int_{-\infty}^{\infty} E(\eta|\xi = x) p_\xi(x) dx \\ &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} y \frac{p(x,y)}{p_\xi(x)} dy \right) p_\xi(x) dx \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y p(x,y) dy dx = E\eta. \end{aligned}$$

When ξ is a discrete random variable, letting $p_i = P(\xi = x_i)$, then

$$E\eta = \sum_i p_i E(\eta|\xi = x_i).$$

This is similar to the total probability formula, called the total expectation formula.

Example

A miner is trapped in a mine containing 3 doors. The first door leads to a tunnel that will take him to safety after 3 hours of travel. The second door leads to a tunnel that will return him to the mine after 5 hours of travel. The third door leads to a tunnel that will return him to the mine after 7 hours. If we assume that the miner is at all times equally likely to choose any one of the doors, what is the expected length of time until he reaches safety?

Solution. Let ξ be the amount of time (in hours) until the miner reaches safety, and let η denote the door he initially chooses.

Solution. Let ξ be the amount of time (in hours) until the miner reaches safety, and let η denote the door he initially chooses. Then

$$E[\xi|\eta = 1] = 3,$$

Solution. Let ξ be the amount of time (in hours) until the miner reaches safety, and let η denote the door he initially chooses. Then

$$E[\xi|\eta=1]=3,$$

$$E[\xi|\eta=2]=5+$$

Solution. Let ξ be the amount of time (in hours) until the miner reaches safety, and let η denote the door he initially chooses. Then

$$E[\xi|\eta=1]=3,$$
$$E[\xi|\eta=2]=5+E\xi,$$

Solution. Let ξ be the amount of time (in hours) until the miner reaches safety, and let η denote the door he initially chooses. Then

$$E[\xi|\eta=1]=3,$$
$$E[\xi|\eta=2]=5+E\xi,$$
$$E[\xi|\eta=3]=7+E\xi.$$

Now

$$
\begin{aligned}
E\xi &= E[\xi|\eta=1]P(\eta=1) \\
&\quad +E[\xi|\eta=2]P(\eta=2) \\
&\quad +E[\xi|\eta=3]P(\eta=3)
\end{aligned}
$$

Now

$$
\begin{aligned}
E\xi &= E[\xi|\eta=1]P(\eta=1) \\
&\quad +E[\xi|\eta=2]P(\eta=2) \\
&\quad +E[\xi|\eta=3]P(\eta=3) \\
&= \frac{1}{3}(3+5+E\xi+7+E\xi) \\
&= 5+\frac{2}{3}E\xi.
\end{aligned}
$$

So

$$E\xi = 15.$$

Example

$\{\xi_i, i \geq 1\}$ i.i.d $\sim B(n,p)$, $\nu \sim P(\lambda)$. ν is independent of $\{\xi_i, i \geq 1\}$. Find $E(\sum_{i=1}^{\nu} \xi_i)$.

Solution.

Example

$\{\xi_i, i \geq 1\}$ i.i.d $\sim B(n,p)$, $\nu \sim P(\lambda)$. ν is independent of $\{\xi_i, i \geq 1\}$. Find $E(\sum_{i=1}^{\nu} \xi_i)$.

Solution. Let $\eta = \sum_{i=1}^{\nu} \xi_i$, then small

$$E(\eta|\nu = r) = E(\sum_{i=1}^{\nu} \xi_i|\nu = r)$$

Example

$\{\xi_i, i \geq 1\}$ i.i.d $\sim B(n,p)$, $\nu \sim P(\lambda)$. ν is independent of $\{\xi_i, i \geq 1\}$. Find $E(\sum_{i=1}^{\nu} \xi_i)$.

Solution. Let $\eta = \sum_{i=1}^{\nu} \xi_i$, then small

$$E(\eta|\nu = r) = E(\sum_{i=1}^{\nu} \xi_i|\nu = r) = E(\sum_{i=1}^{r} \xi_i|\nu = r)$$

Example

$\{\xi_i, i \geq 1\}$ i.i.d $\sim B(n,p)$, $\nu \sim P(\lambda)$. ν is independent of $\{\xi_i, i \geq 1\}$. Find $E(\sum_{i=1}^{\nu} \xi_i)$.

Solution. Let $\eta = \sum_{i=1}^{\nu} \xi_i$, then small

$$
\begin{aligned}
E(\eta|\nu = r) &= E(\sum_{i=1}^{\nu} \xi_i|\nu = r) = E(\sum_{i=1}^{r} \xi_i|\nu = r) \\
&= E(\sum_{i=1}^{r} \xi_i) = \sum_{i=1}^{r} E\xi_i = rnp.
\end{aligned}
$$

So

$$E\eta = \sum_{r=0}^{\infty} E(\eta|\nu = r)P(\nu = r)$$

So

$$\begin{aligned} E\eta &= \sum_{r=0}^{\infty} E(\eta|\nu = r)P(\nu = r) \\ &= np\sum_{r=1}^{\infty} rP(\nu = r) = npE\nu = np\lambda. \end{aligned}$$

Remark Just as conditional probabilities satisfy all of the properties of ordinary probabilities, so do conditional expectations satisfy the properties of ordinary expectations.

Remark Just as conditional probabilities satisfy all of the properties of ordinary probabilities, so do conditional expectations satisfy the properties of ordinary expectations. For instance, the following formulas remain valid.

$$E[g(\eta)|\xi = x] = \sum_j g(y_j) p_{\eta|\xi}(y_j|x)$$

in the discrete case,

$$E[g(\eta)|\xi = x] = \int_{-\infty}^{\infty} g(y) p_{\eta|\xi}(y|x) dy$$

in the continuous case,

Example

The quick-sort algorithm(快速排序法)

Example

The quick-sort algorithm(快速排序法)

设有n个不同的数$x_1, x_2, \ldots, x_n$.我们要将它们按从小到大的次序排列起来$x_{(1)} < x_{(2)} < \ldots < x_{(n)}$. 进行这样的排列需要对这$n$个数两两进行比较，如果全部进行，共需要比较$\frac{n(n-1)}{2}$次，这就是用计算机排列这些数所需要的计算量. 是否有一种算法可以减少比较次数呢?

Example

一种称为快速排序算法(quick-sort algorithm)是这样进行的：从集合$\{x_1, x_2, \ldots, x_n\}$中随机地取一个数$x_J$，将其它数与$x_J$进行比较，把小于$x_J$的数放在其左边，这样的数构成集合$L$，把大于$x_J$的数放在其右边，这样的数构成集合$R$. 然后，对$L$和$R$进行同样处理，依此类推，直到最后每个集合中只有一个数为止. 我们用ξ表示进行比较的总次数,求$q_n = E\xi$.

解: 设用快速排序法排$L(x_J$左边的数)所需要比较进行次数为ξ_L, 排$R(x_J$右边的数)所需要进行比较次数为ξ_R. 则

$$\xi = \xi_L + \xi_R + (n-1).$$

解: 设用快速排序法排$L(x_J$左边的数)所需要比较进行次数为ξ_L, 排$R(x_J$右边的数)所需要进行比较次数为ξ_R. 则

$$\xi = \xi_L + \xi_R + (n-1).$$

在$x_J = x_{(i)}$的条件下, L和R中分别有$i-1$, $n-i$个元素. 所以

$$E[\xi_L | x_J = x_{(i)}] = q_{i-1}, \quad E[\xi_R | x_J = x_{(i)}] = q_{n-i}.$$

解: 设用快速排序法排$L(x_J$左边的数)所需要比较进行次数为ξ_L, 排$R(x_J$右边的数)所需要进行比较次数为ξ_R. 则

$$\xi = \xi_L + \xi_R + (n-1).$$

在$x_J = x_{(i)}$的条件下, L和R中分别有$i-1$, $n-i$个元素. 所以

$$E[\xi_L | x_J = x_{(i)}] = q_{i-1}, \quad E[\xi_R | x_J = x_{(i)}] = q_{n-i}.$$

因此

$$E[\xi | x_J = x_{(i)}] = q_{i-1} + q_{n-i} + (n-1).$$

解: 设用快速排序法排$L(x_J$左边的数)所需要比较进行次数为ξ_L, 排$R(x_J$右边的数)所需要进行比较次数为ξ_R. 则

$$\xi = \xi_L + \xi_R + (n-1).$$

在$x_J = x_{(i)}$的条件下, L和R中分别有$i-1$, $n-i$个元素. 所以

$$E[\xi_L | x_J = x_{(i)}] = q_{i-1}, \quad E[\xi_R | x_J = x_{(i)}] = q_{n-i}.$$

因此

$$E[\xi | x_J = x_{(i)}] = q_{i-1} + q_{n-i} + (n-1).$$

而

$$P(x_J = x_{(i)}) = \frac{1}{n}.$$

所以

$$
\begin{aligned}
q_n =&E\xi = \sum_{i=1}^n E[\xi|x_J = x_{(i)}]P(x_J = x_{(i)}) \\
=&n-1+\frac{1}{n}\sum_{i=1}^n (q_{i-1}+q_{n-i}) = n-1+\frac{2}{n}\sum_{i=1}^n q_{i-1}.
\end{aligned}
$$

所以

$$
\begin{aligned}
q_n =& E\xi = \sum_{i=1}^{n} E[\xi | x_J = x_{(i)}] P(x_J = x_{(i)}) \\
=& n-1+\frac{1}{n}\sum_{i=1}^{n}(q_{i-1}+q_{n-i}) = n-1+\frac{2}{n}\sum_{i=1}^{n} q_{i-1}.
\end{aligned}
$$

$$
nq_n = n(n-1)+2\sum_{i=1}^{n} q_{i-1}.
$$

$$
nq_n-(n-1)q_{n-1} = n(n-1)-(n-1)(n-2)+2q_{n-1}.
$$

$$
nq_n = 2(n-1)+(n+1)q_{n-1}.
$$

$$
\begin{aligned}
\frac{q_n}{n+1} =& \frac{q_{n-1}}{n} + \frac{2(n-1)}{n(n+1)} \\
=& \frac{q_{n-1}}{n} + \frac{2}{n} + 4\left(\frac{1}{n+1} - \frac{1}{n}\right) \\
=& \cdots = 2\sum_{k=1}^{n} \frac{1}{k} + \frac{4}{n+1} - 4 \\
=& 2\left(\log n + \gamma + \frac{1}{2n} + O(n^{-2})\right) - \frac{4n}{n+1}.
\end{aligned}
$$

$$
\begin{aligned}
\frac{q_n}{n+1} =& \frac{q_{n-1}}{n} + \frac{2(n-1)}{n(n+1)} \\
=& \frac{q_{n-1}}{n} + \frac{2}{n} + 4\left(\frac{1}{n+1} - \frac{1}{n}\right) \\
=& \cdots = 2\sum_{k=1}^{n} \frac{1}{k} + \frac{4}{n+1} - 4 \\
=& 2\left(\log n + \gamma + \frac{1}{2n} + O(n^{-2})\right) - \frac{4n}{n+1}.
\end{aligned}
$$

因此

$$
q_n = 2(n+1)\log n + n(2\gamma - 4) + 2\gamma + 1 + O(n^{-1}).
$$