3.1.4 Expectations for functions of random variables

3.1.4 Expectations for functions of random variables

The expectation of a function of a discrete random variable:

ξ has the pmf

$$\begin{pmatrix} x_1 & x_2 & \cdots & x_k & \cdots \\ p_1 & p_2 & \cdots & p_k & \cdots \end{pmatrix}.$$

Then for $\eta = f(\xi)$,

3.1.4 Expectations for functions of random variables

The expectation of a function of a discrete random variable:

ξ has the pmf

$$\begin{pmatrix} x_1 & x_2 & \cdots & x_k & \cdots \\ p_1 & p_2 & \cdots & p_k & \cdots \end{pmatrix}.$$

Then for $\eta = f(\xi)$,

$$\begin{pmatrix} f(x_1) & f(x_2) & \cdots & f(x_k) & \cdots \\ p_1 & p_2 & \cdots & p_k & \cdots \end{pmatrix}.$$

So the pmf of η is

$$\begin{pmatrix} y_1 & y_2 & \cdots & y_i & \cdots \\ p_1^* & p_2^* & \cdots & p_i^* & \cdots \end{pmatrix}$$

where $p_i^* = \sum_{j:f(x_j)=y_i} p_j$.

So the pmf of η is

$$\begin{pmatrix} y_1 & y_2 & \cdots & y_i & \cdots \\ p_1^* & p_2^* & \cdots & p_i^* & \cdots \end{pmatrix}$$

where $p_i^* = \sum_{j:f(x_j)=y_i} p_j$. Hence $E\eta$ exists if and only if

$$\sum_i |y_i| p_i^*$$

So the pmf of η is

$$\begin{pmatrix} y_1 & y_2 & \cdots & y_i & \cdots \\ p_1^* & p_2^* & \cdots & p_i^* & \cdots \end{pmatrix}$$

where $p_i^* = \sum_{j:f(x_j)=y_i} p_j$. Hence $E\eta$ exists if and only if

$$\sum_i |y_i| p_i^* = \sum_i \sum_{j:f(x_j)=y_i} |f(x_j)| p_j$$

So the pmf of η is

$$\begin{pmatrix} y_1 & y_2 & \cdots & y_i & \cdots \\ p_1^* & p_2^* & \cdots & p_i^* & \cdots \end{pmatrix}$$

where $p_i^* = \sum_{j:f(x_j)=y_i} p_j$. Hence $E\eta$ exists if and only if

$$\begin{aligned} \sum_i |y_i| p_i^* &= \sum_i \sum_{j:f(x_j)=y_i} |f(x_j)| p_j \\ &= \sum_k |f(x_k)| p_k < \infty. \end{aligned}$$

Further,

$$E\eta = \sum_i y_i p_i^* = \sum_i \sum_{j:f(x_j)=y_i} f(x_j)p_j = \sum_k f(x_k)p_k.$$

Further,

$$E\eta = \sum_i y_i p_i^* = \sum_i \sum_{j:f(x_j)=y_i} f(x_j)p_j = \sum_k f(x_k)p_k.$$

$$Ef(\xi) = \sum_k f(x_k)p_k = \sum_k f(x_k)P(\xi = x_k).$$

Theorem Suppose that ξ is a discrete random variable with the distribution $F_\xi(x)$ and

$$P(\xi = x_k) = p_k, \quad k = 1, 2, \cdots ,$$

$f(x)$ a Borel function on the real line. Let $\eta = f(\xi)$. Then $Ef(\xi)$ exists if and only if

$$\int_{-\infty}^{\infty} |f(x)| dF_\xi(x) = \sum_k |f(x_k)| P(\xi = x_k) < \infty$$

and

$$Ef(\xi) = \sum_k f(x_k) P(\xi = x_k) = \int_{-\infty}^{+\infty} f(x) dF_\xi(x).$$

In general, we have

Theorem 1. Suppose that ξ is a random variable with the distribution $F_\xi(x)$, $f(x)$ a Borel function on the real line. Let $\eta = f(\xi)$. Then

$$Ef(\xi) = \int_{-\infty}^{+\infty} y dF_\eta(y) = \int_{-\infty}^{+\infty} f(x) dF_\xi(x).$$

In general, we have

Theorem 1. Suppose that ξ is a random variable with the distribution $F_\xi(x)$, $f(x)$ a Borel function on the real line. Let $\eta = f(\xi)$. Then

$$Ef(\xi) = \int_{-\infty}^{+\infty} y dF_\eta(y) = \int_{-\infty}^{+\infty} f(x) dF_\xi(x).$$

When ξ has density $p(x)$, then

$$Ef(\xi) = \int_{-\infty}^{+\infty} f(x)p(x)dx.$$

When ξ is a random variable of the general type and $f(x)$ is a general Borel function, the proof of this theorem is due to measure theory. Next, we give a proof for the case when ξ is a continuous type random variable. Let $p(x)$ be the pdf of ξ. Then

When ξ is a random variable of the general type and $f(x)$ is a general Borel function, the proof of this theorem is due to measure theory. Next, we give a proof for the case when ξ is a continuous type random variable. Let $p(x)$ be the pdf of ξ. Then

$$\begin{aligned} E|\eta| &= \int_0^\infty P(|\eta| > y)dy = \int_0^\infty P(|f(\xi)| > y)dy \\ &= \int_0^\infty \int_{x:|f(x)|>y} p(x)dxdy \\ &= \int_{-\infty}^\infty \int_{y:|f(x)|>y,y\geq 0} p(x)dydx \\ &= \int_{-\infty}^\infty |f(x)|p(x)dx. \end{aligned}$$

So $Ef(\xi)$ exists if and only if $\int_{-\infty}^\infty |f(x)|p(x)dx < \infty$.

Further,

$$
\begin{aligned}
E\eta &= \int_0^{\infty} P(\eta > y)dy - \int_0^{\infty} P(-\eta > y)dy \\
&= \int_0^{\infty} \int_{x: f(x) > y} p(x)dxdy - \int_0^{\infty} \int_{x: -f(x) > y} p(x)dxdy \\
&= \int_{-\infty}^{\infty} \left[\int_{y: f(x) > y, y \geq 0} dy - \int_{y: -f(x) > y, y \geq 0} dy \right] p(x)dx \\
&= \int_{-\infty}^{\infty} f(x)p(x)dx.
\end{aligned}
$$

Further,

$$\begin{aligned}
E\eta &= \int_0^\infty P(\eta > y)dy - \int_0^\infty P(-\eta > y)dy \\
&= \int_0^\infty \int_{x:f(x)>y} p(x)dxdy - \int_0^\infty \int_{x:-f(x)>y} p(x)dxdy \\
&= \int_{-\infty}^\infty \left[\int_{y:f(x)>y,y\geq 0} dy - \int_{y:-f(x)>y,y\geq 0} dy\right] p(x)dx \\
&= \int_{-\infty}^\infty f(x)p(x)dx.
\end{aligned}$$

Theorem 1 tells us that if ξ and η have the same distribution function, then

$$Ef(\xi) = Ef(\eta).$$

On the contrary, if the above equality holds for any bounded continuous function f, then ξ and η have the same distribution function.

In fact, for any z and $\epsilon > 0$, let $f(x)$ be a continuous function such that $f(x) = 1$, $0 \leq f(x) \leq 1$ and $f(x) = 0$ on $(-\infty, z]$, $(z, z + \epsilon]$ and $(z + \epsilon, \infty)$, respectively. Then

In fact, for any z and $\epsilon > 0$, let $f(x)$ be a continuous function such that $f(x) = 1$, $0 \leq f(x) \leq 1$ and $f(x) = 0$ on $(-\infty, z]$, $(z, z+\epsilon]$ and $(z+\epsilon, \infty)$, respectively. Then

$$\begin{aligned} F_\xi(z) &= \int_{-\infty}^{z} f(x) dF_\xi(x) \leq \int_{-\infty}^{\infty} f(x) dF_\xi(x) \\ &= \int_{-\infty}^{\infty} f(x) dF_\eta(x) = \int_{-\infty}^{z+\epsilon} f(x) dF_\eta(x) \\ &\leq \int_{-\infty}^{z+\epsilon} dF_\eta(x) = F_\eta(z+\epsilon). \end{aligned}$$

Letting $\epsilon \to 0$ yields $F_\xi(z) \leq F_\eta(z)$.

In fact, for any z and $\epsilon > 0$, let $f(x)$ be a continuous function such that $f(x) = 1$, $0 \leq f(x) \leq 1$ and $f(x) = 0$ on $(-\infty, z]$, $(z, z+\epsilon]$ and $(z+\epsilon, \infty)$, respectively. Then

$$\begin{aligned} F_\xi(z) &= \int_{-\infty}^{z} f(x)dF_\xi(x) \leq \int_{-\infty}^{\infty} f(x)dF_\xi(x) \\ &= \int_{-\infty}^{\infty} f(x)dF_\eta(x) = \int_{-\infty}^{z+\epsilon} f(x)dF_\eta(x) \\ &\leq \int_{-\infty}^{z+\epsilon} dF_\eta(x) = F_\eta(z+\epsilon). \end{aligned}$$

Letting $\epsilon \to 0$ yields $F_\xi(z) \leq F_\eta(z)$. Similarly, $F_\eta(z) \leq F_\xi(z)$. So ξ and η have the same distribution function.

Example. (Stein's Lemma) (i) Let $\xi \sim N(0,1)$, and g be differentiable function satisfying $|g(x)| \leq c_1 e^{c_2|x|}$ and $|g'(x)| \leq c_1 e^{c_2|x|}$ for some $c_1 > 0, c_2 > 0$. Prove

$$E\left[g(\xi)\xi\right] = Eg'(\xi).$$

(ii)* On the contrary, if the above equality holds for any bounded continuous function $g(x)$ with bounded derivation, then $\xi \sim N(0,1)$.

Proof. For (i), we have

$$
\begin{aligned}
& E\left[g(\xi)\xi\right] \\
= & \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} g(x)x\exp\left\{-\frac{x^2}{2}\right\}dx
\end{aligned}
$$

Proof. For (i), we have

$$\begin{aligned} & E\left[g(\xi)\xi\right] \\ = & \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} g(x)x\exp\left\{-\frac{x^2}{2}\right\}dx \\ = & \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} g(x)d\left[-\exp\left\{-\frac{x^2}{2}\right\}\right]. \end{aligned}$$

Use integration by parts to get

$$
\begin{aligned}
& E\left[g(\xi)\xi\right] \\
= \; & \frac{1}{\sqrt{2\pi}}\left[-g(x)\exp\left\{-\frac{x^2}{2}\right\}\Big|_{-\infty}^{\infty}\right. \\
& \left. + \int_{-\infty}^{\infty} g'(x)\exp\left\{-\frac{x^2}{2}\right\}dx\right] \\
= \; &
\end{aligned}
$$

Use integration by parts to get

$$
\begin{aligned}
& E\left[g(\xi)\xi\right] \\
= \; & \frac{1}{\sqrt{2\pi}}\left[-g(x)\exp\left\{-\frac{x^2}{2}\right\}\Big|_{-\infty}^{\infty}\right. \\
& \left. + \int_{-\infty}^{\infty} g'(x)\exp\left\{-\frac{x^2}{2}\right\}dx\right] \\
= \; & Eg'(\xi).
\end{aligned}
$$

For (ii), Let $\eta \sim N(0,1)$. It is sufficient to show that $Eh(\xi) = Eh(\eta)$ holds for any bounded continuous function $h(x)$.

For (ii), Let $\eta \sim N(0,1)$. It is sufficient to show that $Eh(\xi) = Eh(\eta)$ holds for any bounded continuous function $h(x)$. Without loss of generality, we assume $0 \leq h(x) \leq 1$. Let $g(x)$ satisfy

$$h(x) - Eh(\eta) = g'(x) - xg(x).$$

The above equation is called Stein's equation.

For (ii), Let $\eta \sim N(0,1)$. It is sufficient to show that $Eh(\xi) = Eh(\eta)$ holds for any bounded continuous function $h(x)$. Without loss of generality, we assume $0 \leq h(x) \leq 1$. Let $g(x)$ satisfy

$$h(x) - Eh(\eta) = g'(x) - xg(x).$$

The above equation is called Stein's equation. It can be verified that

$$g(x) = e^{\frac{x^2}{2}} \int_{-\infty}^{x} [h(u) - Eh(\eta)]\, e^{-\frac{u^2}{2}} du$$

is a solution of the equation, and, both $g(x)$ and $g'(x)$ are bounded continuous function.

So

$$\begin{aligned}
Eh(\xi) &= \int_{-\infty}^{\infty} \left[g'(x) - xg(x) + Eh(\eta)\right] dF_\xi(x) \\
&= \int_{-\infty}^{\infty} g'(x) dF_\xi(x) - \int_{-\infty}^{\infty} xg(x) dF_\xi(x) + Eh(\eta) \int_{-\infty}^{\infty} dF_\xi(x) \\
&= Eg'(\xi) - E[\xi g(\xi)] + Eh(\eta) = Eh(\eta).
\end{aligned}$$

The proof is completed. In the second equality above, the linearity of the Stieltjes integral.

Similarly,

Similarly,

$$E\left[\xi g(\xi-1)\right]=E[\lambda g(\xi)],\ \forall g(\text{bounded})\Leftrightarrow \xi\sim P(\lambda).$$

Similarly,

$$E\left[\xi g(\xi-1)\right]=E[\lambda g(\xi)],\ \forall g(\text{bounded}) \Leftrightarrow \xi \sim P(\lambda).$$

Stein-Chen method.

In general,suppose $(\xi_1, \cdots, \xi_n) \sim F(x_1, \cdots, x_n)$. Also, assume that $g(x_1, \cdots, x_n)$ is a Borel function, then

$$Eg(\xi_1, \cdots, \xi_n) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, \cdots, x_n) dF(x_1, \cdots, x_n).$$

In general,suppose $(\xi_1,\cdots,\xi_n) \sim F(x_1,\cdots,x_n)$. Also, assume that $g(x_1,\cdots,x_n)$ is a Borel function, then

$$Eg(\xi_1,\cdots,\xi_n) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g(x_1,\cdots,x_n)dF(x_1,\cdots,x_n).$$

If $(\xi_1,\xi_2,\ldots,\xi_n)$ has pmf $P(\xi_1 = x_1(i_1),\xi_2 = x_2(i_2),\ldots,\xi_n = x_n(i_n)) = p_{i_1 i_2\cdots i_n}$, then

$$Eg(\xi_1,\xi_2,\ldots,\xi_n) = \sum_{i_1,i_2,\ldots,i_n} g\big(x_1(i_1),x_2(i_2),\ldots,x_n(i_n)\big)p_{i_1 i_2\cdots i_n};$$

If $(\xi_1,\xi_2,\ldots,\xi_n)$ has pdf $p(x_1,x_2,\ldots,x_n)$, then

$$Eg(\xi_1,\xi_2,\ldots,\xi_n)$$
$$= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g(x_1,x_2,\ldots,x_n)p(x_1,x_2,\ldots,x_n)dx_1\cdots dx_n.$$

Here, the multi-variable Stieltjes integral is defined similarly as in the one-variable case. For example

$$\int_{a_1}^{b_1}\cdots\int_{a_n}^{b_n} g(x_1,\ldots,x_n)dF(x_1,\ldots,x_n)$$
$$=\lim\sum_{k_1,\cdots,k_n} g(x_1(k_1),\ldots,x_n(k_n))\Delta F(x_1(k_1),\ldots,x_n(k_n)),$$

where $x_i(1), x_i(2), \ldots$ is a partition of $(a_i, b_i]$, $\Delta F(x_1(k_1),\ldots,x_n(k_n))$ is the probability that $(\xi_1,\ldots,\xi_n)$ falls in $(x_1(k_1), x_1(k_1+1)]\times\cdots\times(x_n(k_n), x_n(k_n+1)]$.

In particular, we have

$$E\xi_i = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_i dF(x_1, \cdots, x_n) = \int_{-\infty}^{\infty} x dF_i(x),$$

where $F_i(x)$ is the distribution function of ξ_i.

For $F(x, y)$ it follows

$$E\xi\eta = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy dF(x, y)$$

and

$$E\xi^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^2 dF(x, y),$$

etc.

Example. Suppose R and Θ are indept., $\Theta \sim U(0, 2\pi)$, $R \sim Rayleigh$. Find $Ee^{R \sin \Theta}$.

Solution.

Example. Suppose R and Θ are indept., $\Theta \sim U(0, 2\pi)$, $R \sim Rayleigh$. Find $Ee^{R\sin\Theta}$.

Solution.

$$Ee^{R\sin\Theta}$$

Example. Suppose R and Θ are indept., $\Theta \sim U(0, 2\pi)$, $R \sim Rayleigh$. Find $Ee^{R\sin\Theta}$.

Solution.

$$Ee^{R\sin\Theta} = \int_0^\infty \int_0^{2\pi} e^{r\sin\theta} re^{-r^2/2} \frac{1}{2\pi} d\theta dr$$

Example. Suppose R and Θ are indept., $\Theta \sim U(0,2\pi)$, $R \sim Rayleigh$. Find $Ee^{R\sin\Theta}$.

Solution.

$$
\begin{aligned}
Ee^{R\sin\Theta} &= \int_0^{\infty}\int_0^{2\pi} e^{r\sin\theta} re^{-r^2/2}\frac{1}{2\pi}d\theta dr \\
&= \iint e^{x}\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}dxdy(\text{极坐标变换})
\end{aligned}
$$

Example. Suppose R and Θ are indept., $\Theta \sim U(0, 2\pi)$, $R \sim Rayleigh$. Find $Ee^{R\sin\Theta}$.

Solution.

$$\begin{aligned} Ee^{R\sin\Theta} &= \int_0^\infty \int_0^{2\pi} e^{r\sin\theta} re^{-r^2/2} \frac{1}{2\pi} d\theta dr \\ &= \iint e^x \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}} dxdy(\text{极坐标变换}) \\ &= \int_{-\infty}^\infty e^x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \end{aligned}$$

Example. Suppose R and Θ are indept., $\Theta \sim U(0, 2\pi)$, $R \sim Rayleigh$. Find $Ee^{R\sin\Theta}$.

Solution.

$$
\begin{aligned}
Ee^{R\sin\Theta} &= \int_0^\infty \int_0^{2\pi} e^{r\sin\theta} re^{-r^2/2} \frac{1}{2\pi} d\theta dr \\
&= \iint e^x \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}} dxdy(\text{极坐标变换}) \\
&= \int_{-\infty}^\infty e^x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\
&= e^{\frac{1}{2}} \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-1)^2}{2}} dx = e^{\frac{1}{2}}.
\end{aligned}
$$

3.1.5 Basic properties of expectations-examples

3.1.5 Basic properties of expectations-examples

1. $a \le \xi \le b \implies a \le E\xi \le b$.

3.1.5 Basic properties of expectations-examples

1. $a \le \xi \le b \implies a \le E\xi \le b.$
2. $E\xi_1, \cdots, E\xi_n$ exist $\implies$

$$E(\sum_{i=1}^{n} c_i\xi_i + b) = \sum_{i=1}^{n} c_i E\xi_i + b.$$

3.1.5 Basic properties of expectations-examples

1. $a \leq \xi \leq b \implies a \leq E\xi \leq b.$
2. $E\xi_1, \cdots, E\xi_n$ exist $\implies$

$$E(\sum_{i=1}^{n} c_i\xi_i + b) = \sum_{i=1}^{n} c_i E\xi_i + b.$$

3. $\xi_1, \cdots, \xi_n$ indept. and expectations exist $\implies$

$$E(\xi_1 \cdots \xi_n) = E\xi_1 \cdots E\xi_n.$$

3.1.5 Basic properties of expectations-examples

1. $a \leq \xi \leq b \implies a \leq E\xi \leq b$.
2. $E\xi_1, \cdots, E\xi_n$ exist $\implies$
$$E(\sum_{i=1}^{n} c_i\xi_i + b) = \sum_{i=1}^{n} c_i E\xi_i + b.$$
3. $\xi_1, \cdots, \xi_n$ indept. and expectations exist $\implies$
$$E(\xi_1 \cdots \xi_n) = E\xi_1 \cdots E\xi_n.$$
4. If $\xi \leq \eta$ and the exceptions of ξ and η exist, then $E\xi \leq E\eta$.

Example 12. Suppose that ξ follows the binomial distribution $B(n, p)$, find $E\xi$.

Example 12. Suppose that ξ follows the binomial distribution $B(n,p)$, find $E\xi$.

Solution. Consider a Bernoulli trial and set $p = P(A)$ and

$$\xi_i = \begin{cases} 1, & A \text{ occurs in the } i\text{-th trial}, \\ 0, & A \text{ does not occur in the } i\text{-th trial}. \end{cases}$$

Example 12. Suppose that ξ follows the binomial distribution $B(n,p)$, find $E\xi$.

Solution. Consider a Bernoulli trial and set $p = P(A)$ and

$$\xi_i = \begin{cases} 1, & A \text{ occurs in the } i\text{-th trial}, \\ 0, & A \text{ does not occur in the } i\text{-th trial}. \end{cases}$$

Thus ξ_i follows 0-1 distribution, $E\xi_i = p$

Example 12. Suppose that ξ follows the binomial distribution $B(n,p)$, find $E\xi$.

Solution. Consider a Bernoulli trial and set $p = P(A)$ and

$$\xi_i = \begin{cases} 1, & A \text{ occurs in the } i\text{-th trial} , \\ 0, & A \text{ does not occur in the } i\text{-th trial} . \end{cases}$$

Thus ξ_i follows 0-1 distribution, $E\xi_i = p$ and $\xi = \sum_{i=1}^{n} \xi_i$.

Example 12. Suppose that ξ follows the binomial distribution $B(n,p)$, find $E\xi$.

Solution. Consider a Bernoulli trial and set $p = P(A)$ and

$$\xi_i = \begin{cases} 1, & A \text{ occurs in the } i\text{-th trial }, \\ 0, & A \text{ does not occur in the } i\text{-th trial }. \end{cases}$$

Thus ξ_i follows 0-1 distribution, $E\xi_i = p$ and $\xi = \sum_{i=1}^{n} \xi_i$. Hence $E\xi = np$.

Example 13. Suppose that

$$P(\xi = m) = \frac{\binom{M}{m}\binom{N-M}{n-m}}{\binom{N}{n}}, \quad m = 0, 1, \cdots, n.$$

$(n \leq M \leq N)$. Find $E\xi$.

Solution. Design a sampling without replacement.

Example 13. Suppose that

$$P(\xi = m) = \frac{\binom{M}{m}\binom{N-M}{n-m}}{\binom{N}{n}}, \quad m = 0, 1, \cdots, n.$$

$(n \le M \le N)$. Find $E\xi$.

Solution. Design a sampling without replacement. Let ξ_i be the number of defective goods in the i-th draw. Then $\xi = \sum_{i=1}^{n} \xi_i$.

Example 13. Suppose that

$$P(\xi = m) = \frac{\binom{M}{m}\binom{N-M}{n-m}}{\binom{N}{n}}, \quad m = 0, 1, \cdots, n.$$

$(n \leq M \leq N)$. Find $E\xi$.

Solution. Design a sampling without replacement. Let ξ_i be the number of defective goods in the i-th draw. Then $\xi = \sum_{i=1}^{n} \xi_i$. It is known that $P(\xi_i = 1) = M/N$, so $E\xi_i = M/N$.

Example 13. Suppose that

$$P(\xi = m) = \frac{\binom{M}{m}\binom{N-M}{n-m}}{\binom{N}{n}}, \quad m = 0, 1, \cdots, n.$$

$(n \le M \le N)$. Find $E\xi$.

Solution. Design a sampling without replacement. Let ξ_i be the number of defective goods in the i-th draw. Then $\xi = \sum_{i=1}^{n} \xi_i$. It is known that $P(\xi_i = 1) = M/N$, so $E\xi_i = M/N$. Hence

$$E\xi = \sum_{i=1}^{n} E\xi_i = \frac{nM}{N}.$$

Example 14. Suppose that $\xi_1, \cdots, \xi_n$ are independent identically distributed positive random variables with a common density function $f(x)$. Show for any $1 \leq k \leq n$,

$$E\frac{\xi_1 + \cdots + \xi_k}{\xi_1 + \cdots + \xi_n} = \frac{k}{n}.$$

Proof.

Proof. Notice that $\frac{\xi_k}{\xi_1+\cdots+\xi_n}$ is positive and

$$
\begin{aligned}
& E\frac{\xi_k}{\xi_1+\cdots+\xi_n} \\
= & \int_0^\infty \cdots \int_0^\infty \frac{x_k}{x_1+\ldots+x_n} f(x_1)\cdots f(x_n)dx_1\cdots dx_n
\end{aligned}
$$

Proof. Notice that $\frac{\xi_k}{\xi_1+\cdots+\xi_n}$ is positive and

$$\begin{aligned} & E\frac{\xi_k}{\xi_1+\cdots+\xi_n} \\ = & \int_0^\infty \cdots \int_0^\infty \frac{x_k}{x_1+\ldots+x_n} f(x_1)\cdots f(x_n)dx_1\cdots dx_n \\ = & \int_0^\infty \cdots \int_0^\infty \frac{y_1}{y_1+\ldots+y_n} f(y_1)\cdots f(y_n)dy_1\cdots dy_n \end{aligned}$$

Proof. Notice that $\frac{\xi_k}{\xi_1+\cdots+\xi_n}$ is positive and

$$
\begin{aligned}
& E\frac{\xi_k}{\xi_1+\cdots+\xi_n} \\
= & \int_0^\infty \cdots \int_0^\infty \frac{x_k}{x_1+\ldots+x_n} f(x_1)\cdots f(x_n)dx_1\cdots dx_n \\
= & \int_0^\infty \cdots \int_0^\infty \frac{y_1}{y_1+\ldots+y_n} f(y_1)\cdots f(y_n)dy_1\cdots dy_n \\
= & E\frac{\xi_1}{\xi_1+\cdots+\xi_n}.
\end{aligned}
$$

On the other hand,

$$E\frac{\xi_1}{\xi_1+\cdots+\xi_n}+\cdots+E\frac{\xi_n}{\xi_1+\cdots+\xi_n}$$

On the other hand,

$$\begin{aligned}
& E\frac{\xi_1}{\xi_1+\cdots+\xi_n}+\cdots+E\frac{\xi_n}{\xi_1+\cdots+\xi_n} \\
= \ & E\frac{\xi_1+\cdots+\xi_n}{\xi_1+\ldots+\xi_n}=1.
\end{aligned}$$

On the other hand,

$$
\begin{aligned}
& E\frac{\xi_1}{\xi_1+\cdots+\xi_n}+\cdots+E\frac{\xi_n}{\xi_1+\cdots+\xi_n} \\
= \quad & E\frac{\xi_1+\cdots+\xi_n}{\xi_1+\ldots+\xi_n}=1.
\end{aligned}
$$

It follows that

$$
E\frac{\xi_1}{\xi_1+\cdots+\xi_n}=\ldots=E\frac{\xi_n}{\xi_1+\cdots+\xi_n}=\frac{1}{n}.
$$

On the other hand,

$$\begin{aligned} & E\frac{\xi_1}{\xi_1+\cdots+\xi_n}+\cdots+E\frac{\xi_n}{\xi_1+\cdots+\xi_n} \\ = \quad & E\frac{\xi_1+\cdots+\xi_n}{\xi_1+\ldots+\xi_n}=1. \end{aligned}$$

It follows that

$$E\frac{\xi_1}{\xi_1+\cdots+\xi_n}=\ldots=E\frac{\xi_n}{\xi_1+\cdots+\xi_n}=\frac{1}{n}.$$

Hence

$$E\frac{\xi_1+\cdots+\xi_k}{\xi_1+\cdots+\xi_n}$$

On the other hand,

$$
\begin{aligned}
& E\frac{\xi_1}{\xi_1+\cdots+\xi_n}+\cdots+E\frac{\xi_n}{\xi_1+\cdots+\xi_n} \\
= \quad & E\frac{\xi_1+\cdots+\xi_n}{\xi_1+\ldots+\xi_n}=1.
\end{aligned}
$$

It follows that

$$
E\frac{\xi_1}{\xi_1+\cdots+\xi_n}=\ldots=E\frac{\xi_n}{\xi_1+\cdots+\xi_n}=\frac{1}{n}.
$$

Hence

$$
\begin{aligned}
& E\frac{\xi_1+\cdots+\xi_k}{\xi_1+\cdots+\xi_n} \\
= \quad & E\frac{\xi_1}{\xi_1+\cdots+\xi_n}+\cdots+E\frac{\xi_k}{\xi_1+\cdots+\xi_n}=\frac{k}{n}.
\end{aligned}
$$

Example

A grove of 52 trees is arranged in a circular fashion. If a total of 15 chipmunks (花栗鼠) live in these tress, show that there is a group of 7 consecutive trees that together house at least 3 chipmunks.

Example

A grove of 52 trees is arranged in a circular fashion. If a total of 15 chipmunks (花栗鼠) live in these tress, show that there is a group of 7 consecutive trees that together house at least 3 chipmunks.

解: 给定树j,记它连同它旁边按顺时针方向排列的6颗构成一个邻域U_j, 生活在U_j中的chipmunks个数记为Y_j.

Example

A grove of 52 trees is arranged in a circular fashion. If a total of 15 chipmunks (花栗鼠) live in these tress, show that there is a group of 7 consecutive trees that together house at least 3 chipmunks.

解: 给定树j,记它连同它旁边按顺时针方向排列的6颗构成一个邻域U_j, 生活在U_j中的chipmunks个数记为Y_j. 我们只要证明$EY_j > 2$ $\forall j$, 就说明了存在一个树,使得至少有3个chipmunks 生活在此树的邻域中.

事实上, 如果$Y_j \le 2\ \forall j$, 那么$\sum_{j=1}^{52} Y_j \le 2 * 52$.

事实上, 如果$Y_j \leq 2\ \forall j$, 那么$\sum_{j=1}^{52} Y_j \leq 2 * 52$. 从而

$$E\left[\sum_{j=1}^{52} Y_j\right] \leq 2 * 52.$$

事实上, 如果$Y_j \le 2\ \forall j$, 那么$\sum_{j=1}^{52} Y_j \le 2 * 52$. 从而

$$E\left[\sum_{j=1}^{52} Y_j\right] \le 2 * 52.$$

这与

$$\sum_{j=1}^{52} EY_j > 2 * 52$$

矛盾.

下面求EY_j.

$$令\ \ X_i = \begin{cases} 1, & \text{if chipmunk i live in } U_j, \\ 0, & \text{otherwise.} \end{cases}$$

则$Y_j = \sum_{i=1}^{15} X_i$.

下面求EY_j.

$$令\ \ X_i=\begin{cases}1, & \text{if chipmunk i live in } U_j,\\ 0, & \text{otherwise.}\end{cases}$$

则$Y_j=\sum_{i=1}^{15}X_i$. 因为

$$EX_i=P(X_i=1)=\frac{7}{52}.$$

下面求EY_j.

$$令\ \ X_i = \begin{cases} 1, & \text{if chipmunk i live in } U_j, \\ 0, & \text{otherwise.} \end{cases}$$

则$Y_j = \sum_{i=1}^{15} X_i$. 因为

$$EX_i = P(X_i = 1) = \frac{7}{52}.$$

$$所以EY_j = \sum_{i=1}^{15} EX_i = \frac{105}{52} > 2.$$

Example 15. Make a census on some kind of disease in a community with large population. Now check blood for N citizens in two ways: (1) each person each time, so need check N times. (2) check the mixture of bloods of a group of k people. If the outcome reports no virus, that means all these k people are not of this disease; while if the outcome reports virus, then each person from this group is checked again, so k people need check $k+1$ times in this way. Which way may decrease the number of checks?

Solution. Consider the second way. Denote by ξ the number of times each person needs check in a group of k people in the second way. Then

$$\xi = \begin{cases} 1/k, & \text{none of } k \text{ people is sick} \\ (k+1)/k, & \text{at least one of } k \text{ people is sick.} \end{cases}$$

So

Solution. Consider the second way. Denote by ξ the number of times each person needs check in a group of k people in the second way. Then

$$\xi = \begin{cases} 1/k, & \text{none of } k \text{ people is sick} \\ (k+1)/k, & \text{at least one of } k \text{ people is sick.} \end{cases}$$

So $P(\xi = \frac{1}{k}) = (1-p)^k$,
$P(\xi = 1 + \frac{1}{k}) = 1 - (1-p)^k$.

Solution. Consider the second way. Denote by ξ the number of times each person needs check in a group of k people in the second way. Then

$$\xi = \begin{cases} 1/k, & \text{none of } k \text{ people is sick} \\ (k+1)/k, & \text{at least one of } k \text{ people is sick.} \end{cases}$$

So $P(\xi = \frac{1}{k}) = (1-p)^k$,
$P(\xi = 1 + \frac{1}{k}) = 1 - (1-p)^k$. Hence

$$\begin{aligned} E\xi &= \frac{1}{k}(1-p)^k + (1+\frac{1}{k})(1-(1-p)^k) \\ &= 1-(1-p)^k + \frac{1}{k}. \end{aligned}$$

Basic properties of expectations (continue)

Corollary

Suppose $|\xi| \leq \eta$, $E\eta < \infty$. *Then* $E\xi$ *exists and* $|E\xi| \leq E|\xi| \leq E\eta$.

Basic properties of expectations (continue)

Corollary

Suppose $|\xi| \leq \eta$, $E\eta < \infty$. *Then* $E\xi$ *exists and* $|E\xi| \leq E|\xi| \leq E\eta$.

Proof. For $M > 0$, let $\xi_M = |\xi|$ if $|\xi| \leq M$, and 0 if $|\xi| > M$. Then $0 \leq \xi_M \leq M$. By Property 1, $0 \leq E\xi_M \leq M$.

Basic properties of expectations (continue)

Corollary

Suppose $|\xi| \le \eta$, $E\eta < \infty$. *Then* $E\xi$ *exists and* $|E\xi| \le E|\xi| \le E\eta$.

Proof. For $M > 0$, let $\xi_M = |\xi|$ if $|\xi| \le M$, and 0 if $|\xi| > M$. Then $0 \le \xi_M \le M$. By Property 1, $0 \le E\xi_M \le M$. So $E\xi_M$, $E\eta$ exist, and $\xi_M \le \eta$. It follows that

$$\int_{-M}^{M} |x| dF_\xi(x) = E\xi_M \le E\eta$$

by Property 4.

Basic properties of expectations (continue)

Corollary

Suppose $|\xi| \leq \eta$, $E\eta < \infty$. *Then* $E\xi$ *exists and* $|E\xi| \leq E|\xi| \leq E\eta$.

Proof. For $M > 0$, let $\xi_M = |\xi|$ if $|\xi| \leq M$, and 0 if $|\xi| > M$. Then $0 \leq \xi_M \leq M$. By Property 1, $0 \leq E\xi_M \leq M$. So $E\xi_M$, $E\eta$ exist, and $\xi_M \leq \eta$. It follows that

$$\int_{-M}^{M} |x| dF_\xi(x) = E\xi_M \leq E\eta$$

by Property 4. Hence $E|\xi| = \int_{-\infty}^{\infty} |x| dF_\xi(x) \leq E\eta < \infty$.

Basic properties of expectations (continue)

Corollary

Suppose $|\xi| \le \eta$, $E\eta < \infty$. Then $E\xi$ exists and $|E\xi| \le E|\xi| \le E\eta$.

Proof. For $M > 0$, let $\xi_M = |\xi|$ if $|\xi| \le M$, and 0 if $|\xi| > M$. Then $0 \le \xi_M \le M$. By Property 1, $0 \le E\xi_M \le M$. So $E\xi_M$, $E\eta$ exist, and $\xi_M \le \eta$. It follows that

$$\int_{-M}^{M} |x| dF_\xi(x) = E\xi_M \le E\eta$$

by Property 4. Hence $E|\xi| = \int_{-\infty}^{\infty} |x| dF_\xi(x) \le E\eta < \infty$.

Finally, since $-|\xi| \le \xi \le |\xi|$, by Property 4 we have $-E|\xi| \le E\xi \le E|\xi|$. The proof is completed.

Corollary

Let $p > 1$. If $E|\xi|^p$ exists, then $E|\xi|$ exists.

Corollary

Let $p > 1$. If $E|\xi|^p$ exists, then $E|\xi|$ exists.

Proof. Since $|\xi| \leq 1 + |\xi|^p$, $E|\xi| \leq 1 + E|\xi|^p$.

5. **Markov inequality**: If $E|\xi|$ exists, then
$$P(|\xi| \geq \epsilon) \leq \frac{E|\xi|}{\epsilon}, \quad \forall \epsilon > 0.$$

5. Markov inequality: If $E|\xi|$ exists, then

$$P(|\xi| \geq \epsilon) \leq \frac{E|\xi|}{\epsilon}, \quad \forall \epsilon > 0.$$

5. **Markov inequality**: If $E|\xi|$ exists, then

$$P(|\xi| \geq \epsilon) \leq \frac{E|\xi|}{\epsilon}, \quad \forall \epsilon > 0.$$

In fact, let

$$\eta = \begin{cases} 1, & \text{if } |\xi| \geq \epsilon, \\ 0, & \text{for otherwise.} \end{cases} \quad \text{Then } \eta \leq \frac{|\xi|}{\epsilon}.$$

By Property 4, we have

$$P(|\xi| \geq \epsilon) = E\eta \leq E\left[\frac{|\xi|}{\epsilon}\right] = \frac{E|\xi|}{\epsilon}.$$

6. $P(\xi = 0) = 1$ if and only if $E|\xi| = 0$.

6. $P(\xi = 0) = 1$ if and only if $E|\xi| = 0$.

6. $P(\xi = 0) = 1$ if and only if $E|\xi| = 0$.

In fact, the "only if" part is obvious. For the "if" part, by Property 5 we have

$$P(|\xi| \geq \epsilon) = 0 \text{ for all } \epsilon > 0.$$

So $P(|\xi| > 0) = 0$.

Convergence theorems

7. (*Monotone convergence theorem*). If $0 \le \xi_n(\omega) \nearrow \xi(\omega)$, then

$$\lim_{n\to\infty} E\xi_n = E\xi. \qquad (*)$$

If $0 \le \xi_n(\omega) \searrow 0$, and $E\xi_n$s are finite, then

$$\lim_{n\to\infty} E\xi_n = 0.$$

Convergence theorems

7. (*Monotone convergence theorem*). If $0 \leq \xi_n(\omega) \nearrow \xi(\omega)$, then

$$\lim_{n\to\infty} E\xi_n = E\xi. \qquad (*)$$

If $0 \leq \xi_n(\omega) \searrow 0$, and $E\xi_n$s are finite, then

$$\lim_{n\to\infty} E\xi_n = 0.$$

8. (*Dominated convergence theorem*). If $\xi_n(\omega) \to \xi(\omega)$, $|\xi_n| \leq \eta$ and $E\eta < \infty$, then (*) holds.

Convergence theorems

7. (*Monotone convergence theorem*). If $0 \leq \xi_n(\omega) \nearrow \xi(\omega)$, then
$$\lim_{n\to\infty} E\xi_n = E\xi. \qquad (*)$$
If $0 \leq \xi_n(\omega) \searrow 0$, and $E\xi_n$s are finite, then
$$\lim_{n\to\infty} E\xi_n = 0.$$

8. (*Dominated convergence theorem*). If $\xi_n(\omega) \to \xi(\omega)$, $|\xi_n| \leq \eta$ and $E\eta < \infty$, then (*) holds.

9. (*Bounded convergence theorem*). If $\xi_n(\omega) \to \xi(\omega)$ and $|\xi_n| \leq M < \infty$, then (*) holds.

证明: 先证明有界收敛定理.

首先, 已知$|\xi_n| \leq M$, $|\xi| \leq M$. 由性质1, $E\xi_n$, $E\xi$存在, 并且$|E\xi_n - E\xi| \leq E|\xi_n - \xi|$.

证明: 先证明有界收敛定理.

首先, 已知$|\xi_n| \leq M$, $|\xi| \leq M$. 由性质1, $E\xi_n$, $E\xi$存在, 并且$|E\xi_n - E\xi| \leq E|\xi_n - \xi|$. 对任给的$\epsilon > 0$,
记$A_n = \{|\xi_n - \xi| > \epsilon\}$. 在$A_n^c$上有$|\xi_n - \xi| \leq \epsilon$. 而在$A_n$上有$|\xi_n - \xi| \leq 2M$.

证明: 先证明有界收敛定理.

首先, 已知$|\xi_n| \leq M$, $|\xi| \leq M$. 由性质1, $E\xi_n$, $E\xi$存在, 并且$|E\xi_n - E\xi| \leq E|\xi_n - \xi|$. 对任给的$\epsilon > 0$,
记$A_n = \{|\xi_n - \xi| > \epsilon\}$. 在$A_n^c$上有$|\xi_n - \xi| \leq \epsilon$. 而在$A_n$上有$|\xi_n - \xi| \leq 2M$. 所以

$$|\xi_n - \xi| \leq \epsilon + 2MI_{A_n}.$$

因此

$$|E\xi_n - E\xi| \leq E|\xi_n - \xi| \leq \epsilon + 2MP(A_n).$$

另一方面, 由于$\xi_n(\omega) \to \xi(\omega)$, 所以$\lim_{n\to\infty} A_n = \emptyset$.

另一方面, 由于$\xi_n(\omega) \to \xi(\omega)$, 所以$\lim_{n\to\infty} A_n = \emptyset$. 由概率的连续性得

$$\lim_{n\to\infty} P(A_n) = P(\lim_{n\to\infty} A_n) = 0.$$

另一方面，由于$\xi_n(\omega) \to \xi(\omega)$，所以$\lim_{n\to\infty} A_n = \emptyset$. 由概率的连续性得

$$\lim_{n\to\infty} P(A_n) = P(\lim_{n\to\infty} A_n) = 0.$$

所以

$$\limsup_{n\to\infty} |E\xi_n - E\xi| \le \limsup_{n\to\infty} E|\xi_n - \xi| \le \epsilon.$$

由$\epsilon > 0$的任意性得

$$\lim_{n\to\infty} |E\xi_n - E\xi| = \lim_{n\to\infty} E|\xi_n - \xi| = 0.$$

现在证明单调收敛定理. 设$0 \le \xi_n(\omega) \nearrow \xi(\omega)$. 对任意的$M > 0$, 令$\eta_n = \xi_n I\{|\xi_n| \le M\}$, $\eta = \xi I\{|\xi| \le M\}$. 则$\eta_n \le \xi_n$, $\eta \le \xi$,

$$\eta_n(\omega) \to \eta(\omega), \quad \forall \omega$$

并且$0 \le \eta_n \le M$. 由有界收敛定理和数学期望的单调性知,

$$\lim_{n\to\infty} E\xi_n \ge \lim_{n\to\infty} E\eta_n = E\eta = \int_0^M x dF_\xi(x).$$

令$M \to \infty$得

$$\lim_{n\to\infty} E\xi_n \ge E\xi.$$

如果$E\xi = \infty$, 则结论已经得证.

现在证明单调收敛定理. 设$0 \le \xi_n(\omega) \nearrow \xi(\omega)$. 对任意的$M > 0$, 令$\eta_n = \xi_n I\{|\xi_n| \le M\}$, $\eta = \xi I\{|\xi| \le M\}$. 则$\eta_n \le \xi_n$, $\eta \le \xi$,

$$\eta_n(\omega) \to \eta(\omega), \quad \forall \omega$$

并且$0 \le \eta_n \le M$. 由有界收敛定理和数学期望的单调性知,

$$\lim_{n\to\infty} E\xi_n \ge \lim_{n\to\infty} E\eta_n = E\eta = \int_0^M x dF_\xi(x).$$

令$M \to \infty$得

$$\lim_{n\to\infty} E\xi_n \ge E\xi.$$

如果$E\xi = \infty$, 则结论已经得证. 如果$E\xi < \infty$, 则由于$\xi_n \le \xi$, 由单调性得$E\xi_n \le E\xi$. 所以

$$\lim_{n\to\infty} E\xi_n = E\xi.$$

下设$0 \le \xi_n(\omega) \searrow 0$, $E\xi_n$存在, 这时

$$0 \le \xi_1 - \xi_n \nearrow \xi_1.$$

所以

$$E(\xi_1 - \xi_n) \to E\xi_1.$$

所以

$$E\xi_n \to 0.$$

最后证明控制收敛定理. 记

$$\eta_n = \sup_{m \geq n} |\xi_m - \xi|.$$

则$0 \leq \eta_n(\omega) \searrow 0$.

最后证明控制收敛定理. 记

$$\eta_n = \sup_{m\geq n} |\xi_m - \xi|.$$

则$0 \leq \eta_n(\omega) \searrow 0$. 另一方面, 由于$0 \leq \eta_n \leq 2\eta$, 所以$0 \leq E\eta_n \leq 2E\eta < \infty$.

最后证明控制收敛定理. 记

$$\eta_n = \sup_{m \geq n} |\xi_m - \xi|.$$

则$0 \leq \eta_n(\omega) \searrow 0$. 另一方面, 由于$0 \leq \eta_n \leq 2\eta$, 所以$0 \leq E\eta_n \leq 2E\eta < \infty$. 由单调收敛定理,

$$E\eta_n \to 0.$$

而

$$|E\xi_n - E\xi| \leq E|\xi_n - \xi| \leq E\eta_n.$$

因此

$$\lim_{n \to \infty} E\xi_n = E\xi.$$