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

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 $f(x) = 1$, $0 \le f(x) \le 1$ 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 $f(x) = 1$, $0 \le f(x) \le 1$ and $f(x) = 0$ on $(-\infty, z]$, $(z, z+\epsilon]$ and $(z+\epsilon, \infty)$, respectively. Then

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

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

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

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

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 $h(x)$. Suppose $|h(x)| \leq M$, then $0 \leq \frac{h(x)+M}{2M} \leq 1$. So without loss of generality, we assume $0 \leq h(x) \leq 1$.

For (ii), Let $\eta \sim N(0,1)$. It is sufficient to show that $Eh(\xi) = Eh(\eta)$ holds for any bounded continuous function $h(x)$. Suppose $|h(x)| \leq M$, then $0 \leq \frac{h(x)+M}{2M} \leq 1$. So 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 $h(x)$. Suppose $|h(x)| \le M$, then $0 \le \frac{h(x)+M}{2M} \le 1$. So without loss of generality, we assume $0 \le h(x) \le 1$. 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. (verify! Note $g(\infty) = 0$, $\Phi(x) \le \frac{1}{|x|\sqrt{2\pi}} e^{-x^2/2}$, $\Phi(x) \le e^{-x^2/2}$ for all $x < 0$.)

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

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

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

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

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

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

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.

Solution.

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

Solution.

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

Solution.

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

Solution.

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

Solution.

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