Expectation

(1) Definition:

① For simple r.v.'s:

$\cdot X = \sum_1^n a_i I_{A_i}$. $E(X) = \sum_1^n a_i P(A_i)$. $A_i \in \mathcal{A}$. $\sum A_i = \Omega$.

Lemma. It's well-def: If $\sum_1^n a_i I_{A_i} = \sum_1^m b_i I_{B_i}$. $\Omega = \sum_1^n A_i = \sum_1^m B_i$. Then $\sum a_i P(A_i) = \sum b_i P(B_i)$.

Pf: Another partition: $\Omega = \sum_{i,j} A_i \cap B_j$.

② For nonnegative r.v's:

$\cdot$ Since $\exists X_n(w) = \frac{\lfloor 2^n X(w) \rfloor}{2^n} \wedge n \uparrow X(w)$. $X_n \geq 0$.

Define: $E(X) = \lim_n E(X_n)$.

Lemma. It's well-def: If $\exists$ simple r.v.'s. $X_n, Y_n \uparrow X$, $X_n, Y_n \geq 0$. Then: $\lim E(X_n) = \lim E(Y_n)$. exist.

Pf: Note: $E(X_n)$. $E(Y_n)$ are increasing.

Show: $E(Y_k) \leq \lim_n E(X_n)$. Then let $k \to \infty$.

$\therefore \lim E(Y_k) \leq \lim E(X_n)$. by symmetry. ✓.

Set $A_n = \{X_n > Y_k - \varepsilon\}$. $\therefore X_n \geq (Y_k - \varepsilon) I_{A_n}$.

$A_n \uparrow \Omega$. Check: $E(X_n) \geq E((Y_k - \varepsilon) I_{A_n})$. $n \to \infty$.

Limit operation:

· Recall Fatou's Thm:

i) $X_n \geq Y$. a.s. $E(|Y|) < \infty$. Then $\underline{\lim} E(X_n) \geq E(\underline{\lim} X_n)$.

ii) $X_n \leq Y$. a.s. $E(|Y|) < \infty$. Then $\overline{\lim} E(X_n) \leq E(\overline{\lim} X_n)$.

Cor. $\mu(\underline{\lim} A_n) \geq \underline{\lim}\, \mu(A_n)$. $\mu(\overline{\lim} A_n) \leq \overline{\lim}\, \mu(A_n)$.

③ For general r.v.'s:

· Note: $X = X^+ - X^-$. $E(X) = E(X^+) - E(X^-)$.

Prop. $E_\Omega(X) = E_A(X)$. for $p(A) = 1$.

pf: $|E_{A^c}(X)| \leq \max|X|\, p(A^c) \leq \infty \cdot 0 = 0$.

(2) Integration:

① Def: For nondecreasing, right-conti func on $\mathbb{R}^1$. f.

There exits unique measure μ: $\mu(a,b] = f(b) - f(a)$.

Define: $\int g\, df = \int g\, \mu(dx)$. L-S integral associated with f.

Remark: i) $\int_{(a,b]} f\, d\mu \neq \int_{(a,b)} f\, d\mu$. since μ may not be conti. at $x = b$.

ii) R-S integral require: f. g can't be disconti. at same point. But L-S integral needn't it.

② Some cases:

For: $\int_B f\,dG$. $B \in \mathcal{B}$: (L-S integral)

i) G is right-conti BV:

Note: $G = G_1 - G_2$. nondecreasing Func's difference.

$\therefore \int_{(a,b]} f\,dG \stackrel{\Delta}{=} \int_{(a,b]} f\,dG_1 - \int_{(a,b]} f\,dG_2$.

ii) G is discrete:

Suppose $\{x_k\}_1^\infty$ is its jumps. $\Delta G(x_n) = G(x_n) - G(x_n-)$.

Then: $\int_{(s,t]} f\,dG = \sum_{s<x_k\le t} f(x_k)\,\Delta G(x_k)$

iii) G is absolutely conti:

$\exists g$. $g = G'$ a.e. $\mu(s,t] = \int_{(s,t]} dG = \int_{(s,t]} g\,dx$.

Then: $\int_B f\,dG = \int_B fg\,dx$.

iv) G is mixture of ii). iii). right-conti:

Suppose $G(t) = G(a) + \int_a^t g(x)\,dx + \sum_{x_n\le t} \Delta G(x_n)$.

Then: $\int_{(s,t]} f\,dG = \int_{(s,t]} fg\,dx + \sum_{s<x_n\le t} f(x_n)\,\Delta G(x_n)$

③ Integration by part:

Thm. F. G are BV. right-conti. Then. we have:

$$F(t)G(t) - F(s)G(s) = \int_{(s,t]} F(x-)\,dG(x) + G(x)\,dF(x).$$

$$= \int_{(s,t]} F(x)\,dG(x) + G(x-)\,dF(x).$$

Pf: WLOG. set $s=0<t$.

$$[F(t)-F(0)][G(t)-G(0)] = \int_{(0,t]} dF \int_{(0,t]} dG$$

$$= \int_{(0,t]^2} I_{\{0<x<y\}} + I_{\{0<y\le x\}}\, dF(x)\,dG(y)$$

Apply Fubini Thm. with: $\int_{(0,t]^2} I_{\{0<x<y\}}\, dF\,dG$

$$= \int_{(0,t]} F(y-) - F(0)\, dG(y).$$

Cor. $F(t)G(t) - F(s)G(s) = \int_{(s,t]} F(x-)\,dG + G(x-)\,dF + \sum_{s<x_n\le t} \Delta F(x_n)\Delta G(x_n)$

Pf: show: $\int_{(s,t]} \Delta G(x)\,dF = \sum_{s<x_n\le t} \Delta F(x_n)\,\Delta G(x_n)$

Since $F = F_c + F_d$. $\int_{(s,t]} \Delta G\, dF_c = 0$.

by $\Delta G \neq 0$ on countable points. ($\therefore$ null-measure)

(3) Calculation:

① Thm. (Change of variables)

i) X is measurable: $(\Omega, \mathcal{A}, P) \to (\Omega_0, \mathcal{A}_0, P_X)$.

$P_X = P\circ X^{-1}$. induced measure.

ii) g is Borel on $(\Omega_0, \mathcal{A}_0)$. $g \geq 0$ or $E(|g(X)|) < \infty$ holds.

Then. $E(g(X)) = \int_{\Omega_0} g(y)\, P_X(dy)$.

Remark: Note: $E(g(X)) = \int_\Omega g(X(\omega))\, dP(\omega) = \int_{\Omega_0} g(y)\, dP_X(y)$.

We choose $(\Omega_0, \mathcal{A}_0, P_X) = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}, P_X)$ commonly.

Then apply formulars of L-S integral.

Pf: Four steps: $I_A \rightarrow$ Simple Func. $\rightarrow$ Nonnegative

$\rightarrow$ general.

② Application:

i) Absolutely Conti r.v's:

Lemma. Denote $F_X(x) = \int_{-\infty}^{x} f\, dt$. correspond p.m: P_X.

Then $P_X(B) = \int_B f\, dt$. $\forall B \in \mathcal{B}_{\mathbb{R}^1}$.

Pf: $\mathcal{A} = \{A \in \mathcal{B}_{\mathbb{R}^1} \mid P_X(A) = \int_A f\, dt\}$.

$\mathcal{S} = \{(-\infty, x] \mid x \in \mathbb{R}^1\} \subset \mathcal{A}$. σ-algebra.

Thm. $E(g(X)) = \int g(x) f(x)\, dx$. f is density of X.

Pf: prove: $\int g(x)\, P_X(dx) = \int g(x) f(x)\, dx$.

Four steps as usual by Lemma.

ii) Discrete r.v's:

Thm. suppose $P(X=x_k)=p_k$. $E(g(X))=\sum p_k g(x_k)$.

Pf: $g(X)$ is r.v. write: $g(X)=\sum g(x_k) I_{\{x_k\}}$.

1°) $g\geq 0$. Truncate: $g_n=\sum_{k\leq n} g(x_k) I_{\{x_k\}}$.

Apply MCT: $g_n \uparrow g$.

2°) general. $g=g^+-g^-$.

(4) Relation with Tail prob.

① Thm. $\sum_{n=1}^{\infty} P(|X|\geq n) \leq E(|X|) \leq 1+\sum_{n=1}^{\infty} P(|X|\geq n)$.

So $E(|X|)<\infty \Leftrightarrow \sum P(|X|\geq n)<\infty$.

Pf: $A_n=\{n\leq |X|<n+1\}$. $E(|X|)=\sum E_{A_n}(|X|)$

Note: $\sum_1^m nP(A_n)=\sum_1^m P(|X|\geq n)-mP(|X|\geq m+1)$

1°) $E(|X|)<\infty$. $mP(|X|\geq m+1)\leq E(|X| I_{\{|X|\geq m+1\}})\to 0$

$\therefore \sum nP(A_n)=\sum P(|X|\geq n)$.

2°) $E(|X|)=\infty \leq \sum P(|X|\geq n)$. it's trivial.

Cor. If $X:\Omega\longrightarrow \mathbb{Z}$, $w\longmapsto k$. Then $E(|X|)=\sum_{n=1}^{\infty} P(|X|\geq n)$

Pf: $E(|X|)=\sum n P(|X|=n)=\sum P(|X|\geq n)$.

Cor. $E(|X|) < \infty \iff \sum p(|X| \geq cn) < \infty$.

for $\forall c \in \mathbb{R}^+$.

Pf: Apply on $|X|/c$. $E(|X|/c) < \infty \iff E(|X|) < \infty$.

Cor. $\{X, X_n\}$ i.i.d. $\overline{\lim} \frac{|X_n|}{n} \underset{a.s}{\leq} c \iff E(|X|) < \infty$. $\forall c > 0$.

Pf: $\sum p(|X_n| \geq cn) < \infty \iff p(|X_n| \geq cn, i.o) = 0$.

$\iff \overline{\lim} \frac{|X_n|}{n} \leq c$. a.s. $\forall c > 0$

② **Thm.** If $Y \geq 0$. Then. $E(Y) = \int_0^\infty p(Y \geq y) = \int_0^\infty p(Y > y)\, dy$.

Pf: Consider turn Y into integer values:

$Y_n = \frac{[2^n Y]}{2^n} \nearrow Y$. $Y_n \geq 0$. Denote $X_n = 2^n Y_n$

$\therefore E(X_n) = \sum_{k=1} p(X_n \geq k) = \sum p(2^n Y \geq k)$

$A_k^n = [\frac{k}{2^n} \leq y \leq \frac{k+1}{2^n}]$. $\int_0^\infty \square = \sum \int_{A_k^n} p(Y \geq y)\, dy$.

$\therefore E(Y_n) = E(X_n)/2^n \leq \int_0^\infty p(Y \geq y)\, dy \leq \frac{E(X_n)+1}{2^n}$.

By DCT. Let $n \to \infty$. $\therefore E(Y_n) \longrightarrow E(Y)$.

The 2^{nd} "$=$" is from: $\int_0^\infty p(Y = y)\, dy = 0$.

Since $p(Y = y) \neq 0$ only on countable y. (null measure)

Cor. For $Y \geq 0$, $r > 0$. $E(Y^r) = r \int_0^\infty y^{r-1} p(Y \geq y)\, dy$.

$= r \int_0^\infty y^{r-1} p(Y > y)\, dy$.

Pf: $E(Y^r) = \int_0^\infty P(Y^r \geq \eta) d\eta = \int_0^\infty P(Y \geq \eta^{\frac{1}{r}}) d\eta$

$\overset{z=\eta^{\frac{1}{r}}}{=} r\int_0^\infty z^{r-1} P(Y \geq z) dz.$

Cor. $Y \in L^1$. Then $E(Y) = E(Y^+) - E(Y^-)$

$= \int_0^\infty P(Y > \eta) - P(Y \leq -\eta) d\eta.$

Cor. $\forall r > 0$. $E(|X|^r) < \infty \Leftrightarrow \sum n^{r-1} P(|X| \geq n) < \infty.$

Pf: Discretize $\int_0^\infty r x^{r-1} P(|X| \geq x) dx$.

Cor. i) $E(|X|^r) < \infty$. $r > 0 \Rightarrow x^r P(|X| > x) = o(1)$ $(x \to \infty)$.

ii) $x^r P(|X| > x) = o(1) \Rightarrow E(|X|^{r-\varepsilon}) < \infty$. $\forall \varepsilon \in (0, r)$.

But it fails when $\varepsilon = 0$.

Pf: i) $x^r P(|X| \geq x) \leq E(|X|^r I_{\{|X| \geq x\}}) \longrightarrow 0$

ii) $E(|X|^{r-\varepsilon}) < \infty \Leftrightarrow \sum n^{r-1-\varepsilon} P(|X| \geq n) < \infty.$

$n^r P(|X| \geq n) = o(1)$. $\therefore$ $P(|X| \geq n) \sim n^{-r}$.

$\sum n^{-r-1-\varepsilon} P(|X| \geq n) \sim \sum n^{-1-\varepsilon} < \infty.$

Remark: Let $r = 1$. $P(X > x) = \frac{1}{x \ln x}$. Counterexample.

(5) Inequility:

① one-side Markov :

i) $p(X \geq \mu + \lambda) \leq \frac{\sigma^2}{\sigma^2+\lambda^2}$

for $\forall \lambda \geq 0$.

ii) $p(X \leq \mu - \lambda) \leq \frac{\sigma^2}{\sigma^2+\lambda^2}$

<u>Pf</u>: WLOG. $E(X) = 0$. Let $X - \mu = X$.

$$p(X \geq \lambda) = p(X + \lambda \geq \lambda + \lambda) \leq \frac{E((X+\lambda)^2)}{(\lambda+\lambda)^2}$$

$$\left(\frac{E((X+\lambda)^2)}{(\lambda+\lambda)^2}\right)_{min} = \frac{\sigma^2}{\sigma^2+\lambda^2}, \quad \lambda = \sigma^2/\lambda.$$

② <u>Moments of indept. r.v's:</u>

X, Y are indept r.v's.

i) $X + Y \in L^r \Rightarrow X, Y \in L^r$.

ii) $X \in L^p$. for some $p \geq 1$. $E(Y) = 0$. Then:

$$E(|X|^p) \leq E(|X+Y|^p).$$

<u>Pf</u>: i) <u>Lemma</u>. For all large λ:

$$p(|X| \geq \lambda) \leq 2p(|X| \geq \lambda, |Y| \leq \tfrac{\lambda}{2}) \leq 2p(|X+Y| > \lambda/2)$$

<u>Pf</u>: $p(|X| \geq \lambda) = p(|X| \geq \lambda, |Y| \leq \frac{\lambda}{2}) + p(|X| \geq \lambda, |Y| \leq \frac{\lambda}{2})$

$\leq 2p(|X| \geq \lambda, |Y| \leq \frac{\lambda}{2})$, for large λ.

$\leq 2p(|X+Y| \geq \frac{\lambda}{2})$.

$$\Rightarrow E(|X|^r) = \int_0^\infty r x^{r-1} p(|X| \geq x) = \int_0^\lambda + \int_\lambda^\infty \square$$

$$\leq \int_0^r \square + c\int_{2\lambda}^\infty r x^{r-1} p(|X+Y| \geq x)\, dx < \infty.$$

ii) Check: $|1+x|^p \geq 1 + px$. $\forall p \geq 1$, $\forall x \in \mathbb{R}$.

Then $E(|X+Y|^p) = E(|X|^p |1 + Y/X|^p) \geq E(|X|^p) + E(Y) E\left(\frac{p|X|^p}{X}\right)$

$= E(|X|^p)$