9 MATH3733

9.1 More about conditional expectations

- The following formula can be shown by induction, similarly to the proof of lemma 1 from lecture 6.
  $$E(E(X_n|\mathcal{F}^X_{n-1})|\mathcal{F}^X_{n-2}) = E(X_n|\mathcal{F}^X_{n-2}).$$
  More generally, for any $0 < k \le n$, by induction
  $$E(E(X_n|\mathcal{F}^X_{n-1})|\mathcal{F}^X_{n-k}) = E(X_n|\mathcal{F}^X_{n-k}).$$
  They say, while taking conditional expectations, a **smaller** filtration "eats" a greater one. $\mathcal{F}^X_{n-2}$ is smaller than $\mathcal{F}^X_{n-1}$ in the usual sense: at time $n-2$ less information is available than at time $n-1$.

  Exercise 1 *Show it.*

- If X and Y are independent then $E(X|Y) = EX$. **Hint:** $p(x,y) = p_X(x)p_Y(y)$, hence, $p_{X|Y}(x|Y) \equiv p_X(x)$ (continuous case); likewise in discrete case.

9.2 Formula for call in general CRR model

The theorem 1 from lecture 6 for **any** step may be presented as

$$\exp(-r\delta)\tilde{E}(C_{t+\delta}|\mathcal{F}^C_t) = C_t,$$

that is, for any step the process C_t is a discounted martingale with respect to implied probabilities. (The same holds true for S_t; but we are interested in C_t.) Using induction and the remark on conditional expectations above, we may conclude that for any t and any $k > 0$,

$$\exp(-rk\delta)\tilde{E}(C_{t+k\delta}|\mathcal{F}^C_t) = C_t.$$

In particular, with $t = 0$ and $k\delta = T$ this gives us the value of call option in general CRR model:

$$\exp(-rT)\tilde{E}(C_T|\mathcal{F}^C_0) \equiv \exp(-rT)\tilde{E}(C_T) = C_0. \tag{1}$$

9.3 Independent random variables

Definition 1 *We call r.v.'s* $X_1, X_2, \ldots$ **independent** *if either of the following holds:*

1. *For any* n *and any* $x_1, x_s, \dots, x_n$,

$$P(X_1 \le x_1, \dots, X_n \le x_n) = \prod_{k=1}^{n} P(X_k \le x_k);$$

2. *for any* n*, if the random vector* $(X_1, \dots, X_n)$ *is continuous, for any* $x_1, x_s, \dots, x_n$,

$$p_{X_1,\dots,X_n}(x_1, \dots, x_n) = \prod_{k=1}^{n} p_{X_k}(x_k);$$

3. *for any* n *and for any bounded functions* $f_1(x), \dots, f_n(x)$,

$$E\prod_{k=1}^{n} f(X_k) = \prod_{k=1}^{n} Ef(X_k).$$

Concerning equivalence see textbooks on probability.

9.4 Weak convergence and characteristic functions

There are three **equivalent** notions of "weak probabilistic convergence" (again, see textbooks on probability).

Definition 2 *Sequence of random variables* $X_1, X_2, \dots$ **converges weakly** *to a r.v.* X_0*, notation* $X_n \Longrightarrow X_0$*, if any of the following holds true:*

1. $F_{X_n}(x) \to F_X(x)$*, for any* x *where* F_{X_0} *is continuous* $(F_{X_n}(x) = P(X_n \le x))$*;*
2. $Ef(X_n) \to Ef(X)$, $n \to \infty$, $\forall f$ *continuous and bounded;*
3. $E\exp(i\lambda X_n) \to E\exp(i\lambda X)$, $n \to \infty$, $\forall \lambda \in R$.

Definition 3 *An expression* $E\exp(i\lambda X)$ *as function of* $\lambda \in R$ *is called* **characteristic function** *of r.v.* X.

Hence, if we are going to show that (cf. below)

$$\frac{S_n - na}{\sigma\sqrt{n}} \Longrightarrow Y \sim \mathcal{N}(0,1), \tag{2}$$

we may use any of equivalent definitions above. A connection between the second and the third definitions appeals to Fourier analysis, namely, the Weierstrass theorem saying that any continuous function $f(x)$ on $[a,b]$ can be approximated by finite sums of trigonometric polynomials of the form $\sum_k \exp(i\lambda_k x)$. [In turn, this may be shown using simple probabilistic methods, namely, the Law of Large Numbers theorem, see [W.Feller, An Introduction To Probability Theory And Its Applications, vol.1]].

9.5 Characteristic function for $\mathcal{N}(0,1)$

To use the third definition, we may need to know the characteristic function for $\mathcal{N}(0,1)$, which we now calculate:

1. **Exercise 2** *Let $Y \sim \mathcal{N}(0,1)$. Show that for any $n = 1, 3, 5, \ldots$, $EY^n = 0$, while for any $n = 2k,\ k = 1, 2, \ldots$,*

   $$EY^n = (n-1)!! \qquad (k!! = k(k-2)\ldots).$$

   Hint: integrate by parts using $\int(x\exp(-x^2/2))dx = -\exp(-x^2/2)$, and use induction.

2. $$Ee^{i\lambda Y} = E\sum_{n=0}^{\infty}\frac{(i\lambda Y)^n}{n!} = \sum_{n=0}^{\infty}E\frac{(i\lambda Y)^n}{n!} = \sum_{k=0}^{\infty}E\frac{(i\lambda Y)^{2k}}{(2k)!} = \sum_{n=0}^{\infty}\frac{(-\lambda^2)^k((2k-1)!!)}{(2k)!}$$

   $$= \sum_{n=0}^{\infty}\frac{(-\lambda^2)^k(2k-1)(2k-3)\ldots}{(2k)(2k-1)(2k-2)\ldots} = \sum_{n=0}^{\infty}\frac{(-\lambda^2)^k}{(2k)(2k-2)\ldots} = \sum_{n=0}^{\infty}\frac{(-\lambda^2)^k}{2^k\,k!} = e^{-\lambda^2/2}.$$

9.6 Properties of characteristic functions

Lemma 1 *Assuming that all the expectations below exist,*

$$(E\exp(i\lambda X))' = E(iX)\exp(i\lambda X)), \qquad E\exp(i\lambda X))'' = -E(X)^2\exp(i\lambda X)),$$
$$E\exp(i\lambda X))''' = -iE(X)^3\exp(i\lambda X)).$$

Proof. We will show the formulas for finite sums in discrete case:

$$(E\exp(i\lambda X))' = (\sum_k \exp(i\lambda x_k)p_k)' = \sum_k(\exp(i\lambda x_k)p_k)'$$

$$= \sum_k(ix_k)\exp(i\lambda x_k)p_k = E(iX)\exp(i\lambda X),$$

and similarly,

$$(E\exp(i\lambda X))'' = \sum_k(\exp(i\lambda x_k)p_k)'' = \sum_k(ix_k)^2\exp(i\lambda x_k)p_k = -E(X)^2\exp(i\lambda X),$$
$$E\exp(i\lambda X))''' = \sum_k(\exp(i\lambda x_k)p_k)''' = \sum_k(ix_k)^3\exp(i\lambda x_k)p_k = -iE(X)^3\exp(i\lambda X).$$

Lemma 2 *Let $EX = a$, $EX^2 = b$, $E|X|^3 = c < \infty$. Then*

$$E\exp(i\lambda X) = 1 + i\lambda a - \lambda^2 b/2 + o(\lambda^2)\ (\lambda \to 0), \ \ \textit{where } o(\lambda^2) = O(c^3|\lambda|^3). \qquad (3)$$

Proof follows from the **Taylor expansion** with two terms and the reminder in the integral or Lagrange form, if we use the previous lemma and assertions

$(E\exp(i\lambda X))'|_{\lambda=0} = iE(X)$, $(E\exp(i\lambda X))''|_{\lambda=0} = i^2E(X)^2$, $(E\exp(i\lambda X))'''|_{\lambda=0} = i^3E(X)^3$.

9.7 Central Limit Theorem

(The following theorem can be proved without assumption $E|X|^3 < \infty$, but we would like to simplify the calculus and prepare ourselves to theorem 2 below.)

Theorem 1 *Let* $X_1, X_2, \ldots$ *be independent and identically distributed r.v.'s, such that* $EX_k = a$, $var(X_k) = \sigma^2 > 0$, *and* $c = E|X_k|^3 < \infty$; *denote* $S_n = \sum_{k=1}^n X_k$. *Then*

$$\frac{S_n - na}{\sigma\sqrt{n}} \Longrightarrow Y \sim \mathcal{N}(0,1).$$

Proof. We use characteristic functions:

$$E\exp(\frac{i\lambda(S_n - na)}{\sigma\sqrt{n}}) = \prod_{k=1}^n E\exp(\frac{i\lambda(X_k - a)}{\sigma\sqrt{n}}) \quad \text{(we used independence)}$$

$$= \left(E\exp(\frac{i\lambda(X_1 - a)}{\sigma\sqrt{n}})\right)^n \quad \text{(we used that all } X\text{'s are identically distributed)}$$

$$= \left(1 - \frac{\lambda^2\sigma^2}{2n\sigma^2} + O(\frac{|\lambda|^3 c}{\sigma^3 n^{3/2}})\right)^n \quad \text{(we used lemma 2)}$$

$$\to \exp(-\lambda^2/2), \quad n \to \infty. \quad \text{The latter is the characteristic function of } Y.$$

9.8 Another CLT

There is a lot of versions of CLT. In our course we will need the following one.

Theorem 2 *Let for any n,* $X_1(n), X_2(n), \ldots$ *be independent and identically distributed r.v.'s, – perhaps different for different n's, – such that* $EX_k(n) = a(n)$, $var(X_k) = \sigma^2(n) \to \sigma^2 > 0$, *and* $c_n = E|X_k|^3 \le c \le \infty$; *denote* $S_n = \sum_{k=1}^n X_k(n)$. *Then*

$$\frac{S_n - na(n)}{\sigma(n)\sqrt{n}} \Longrightarrow Y \sim \mathcal{N}(0,1).$$

Proof, in fact, repeats the previous proof: as $n \to \infty$,

$$E\exp(\frac{i\lambda(S_n(n) - na(n))}{\sigma(n)\sqrt{n}}) = \prod_{k=1}^n E\exp(\frac{i\lambda(X_k - a(n))}{\sigma(n)\sqrt{n}}) = \left(E\exp(\frac{i\lambda(X_1 - a(n))}{\sigma(n)\sqrt{n}})\right)^n$$

$$= \left(1 - \frac{\lambda^2\sigma^2(n)}{2n\sigma^2(n)} + O(\frac{|\lambda|^3 c_n}{\sigma(n)^3 n^{3/2}})\right)^n \sim \left(1 - \frac{\lambda^2}{2n} + O(\frac{|\lambda|^3 c}{\sigma^3 n^{3/2}})\right)^n \to \exp(-\lambda^2/2).$$