Lectures 20-21

Now we state the following theorem which provides a useful tool to compute expection of random variable which are mixed in nature.

Theorem 0.1 *Let X be a continuous random variable with pdf f and $\phi : \mathbb{R} \to \mathbb{R}$ be a continuous function such that the integral $\int_{-\infty}^{\infty} \phi(x)f(x)dx$ is finite. Then*

$$E[\phi \circ X] \;=\; \int_{-\infty}^{\infty} \phi(x)f(x)dx\,.$$

Proof:(outline) First, I will give a proof for a special case, i.e. when $\varphi : \mathbb{R} \to \mathbb{R}$ is strictly monotone and differentiable. Set $Y = \varphi(X)$. Then Y has density g given by

$$g(y) = \; f(\varphi^{-1}(y))|(\varphi^{-1})'(y)|, y \in \varphi(\mathbb{R}), \; = 0 \;\; \text{otherwise.}$$

i.e.

$$g(y) = f(\varphi^{-1}(y))\Big|\frac{1}{\varphi'(\varphi^{-1}(y))}\Big|, y \in \varphi(\mathbb{R}), \; = 0 \;\; \text{otherwise.}$$

Hence

$$\begin{aligned}
E[\varphi(X)] &= \int_{-\infty}^{\infty} yg(y)dy \\
&= \int_{\varphi(\mathbb{R})} yf(\varphi^{-1}(y))\Big|\frac{1}{\varphi'(\varphi^{-1}(y))}\Big|dy \\
\text{(use } y = \varphi(x),\ \text{Jacobian is } \frac{dy}{dx} = \varphi'(x)) &= \int_{-\infty}^{\infty} \varphi(x)f(x)\Big|\frac{1}{\varphi'(x)}\varphi'(x)\Big|dx \\
&= \int_{-\infty}^{\infty} \varphi(x)f(x)dx.
\end{aligned}$$

In particular, when $\varphi(x) = x^{2n+1}, n \geq 0$ and $E[X^{2n+1}]$ exists. Then note that φ is (strictly) increasing and differentiable. Hence

$$E[X^{2n+1}] \;=\; \int_{-\infty}^{\infty} x^{2n+1}f(x)dx.$$

When $\varphi(x) = x^{2n}$. Then note that φ is not one to one.

Set $Y = X^{2n}$ and G, g denote respectively the distribution function and the pdf of Y. Then for $y > 0$,

$$\begin{aligned}
G(y) &= P\big(Y \leq y\big) \\
&= P\big(-y^{\frac{1}{2n}} \leq X \leq y^{\frac{1}{2n}}\big) \\
&= F(y^{\frac{1}{2n}}) - F(-y^{\frac{1}{2n}}).
\end{aligned}$$

Hence

$$g(y) \;=\; \frac{1}{2n} y^{\frac{1}{2n}-1}\Big(f(y^{\frac{1}{2n}}) + f(-y^{\frac{1}{2n}})\Big), y > 0, \;=\; 0 \text{ for } y \le 0.$$

Therefore

$$\begin{aligned}
E[X^{2n}] &= \int_{-\infty}^{\infty} y g(y) dy \\
&= \frac{1}{2n}\int_0^{\infty} y^{\frac{1}{2n}}\Big(f(y^{\frac{1}{2n}}) + f(-y^{\frac{1}{2n}})\Big)dy \\
&= \frac{1}{2n}\int_0^{\infty} y^{\frac{1}{2n}} f(y^{\frac{1}{2n}})dy + \frac{1}{2n}\int_0^{\infty} y^{\frac{1}{2n}} f(-y^{\frac{1}{2n}})dy \\
(\text{use } y = x^{2n}) &= \int_0^{\infty} x f(x) x^{2n-1} dx + \int_0^{\infty} x f(-x) x^{2n-1} dx \\
(\text{use } u = -x) &= \int_0^{\infty} x^{2n} f(x) dx + \int_{-\infty}^{0} u^{2n} f(u) du \\
&= \int_{-\infty}^{\infty} x^{2n} f(x) dx.
\end{aligned}$$

When $\varphi(x) = a_0 + a_1 x + \cdots + a_n x^n$, a polynomial of degree n, then using Theorem 0.3, Lecturenotes18-19, we get

$$\begin{aligned}
E\varphi(X) &= \sum_{k=0}^{n} a_k E X^k \\
&= \sum_{k=0}^{n} a_k \int_{-\infty}^{\infty} x^k f(x) dx \\
&= \int_{-\infty}^{\infty} \varphi(x) f(x) dx.
\end{aligned}$$

Now suppose X is taking values in $[a,\ b]$ for some $-\infty < a < b < \infty$, the pdf f is continuous and $\varphi : [a,\ b] \to \mathbb{R}$ is continuous.

In this case, we use Weierstrass approximation theorem (i.e. there exists a sequence of polynomials p_n such that $p_n(x) \to \varphi(x)$ uniformly in $x \in [a,\ b]$.

Then $p_n(X_n(\omega) \to \varphi(X(\omega)) \forall \omega$ (in fact the convergence is uniform in ω. Also

$$|p_n(X)| \le K,$$

where $K = \sup_n \max\{p_n(x) : x \in [a,\ b]\}$. Hence using dominated convergence theorem,

$$\lim_{n\to\infty} Ep_n(X) = E\varphi(X).$$

Also using uniform convergence, it follows that

$$\lim_{n\to\infty} p_n(x)f(x)dx = \int \varphi(x)f(x)dx.$$

Now from

$$Ep_n(X) = \int p_n(x)f(x)dx$$

and the above limits, we get

$$E\varphi(X) = \int \varphi(x)f(x)dx$$

Now slightly more refined use of dominated convergence theorem, we can see the above result for any continuous φ and any X with a pdf. $\square$

Example 0.1 *Let* $X \sim U(0,2)$ *and* $Y = \max\{1, X\}$*. Then for* $\varphi(x) = \max\{1, x\}$*, we have*

$$\begin{aligned} EY &= E[\varphi \circ X] \\ &= \int_{-\infty}^{\infty} \varphi(x)f(x)dx \\ &= \frac{1}{2}\int_0^2 \max\{1, x\}dx \\ &= \frac{1}{2} + \frac{1}{2}\int_1^2 xdx \\ &= \frac{5}{4}, \end{aligned}$$

where f *is the pdf of* $U(0,2)$*.*

Example 0.2 *Let* X *be a random variable with pdf* f*. Find* $E[XI_{\{X\leq a\}}]$ *where* $a \in \mathbb{R}$*.*

Note this doesn't come under the realm of Theorem 0.1, because here $\varphi(x) = xI_{(-\infty,\, a]}(x)$ *has a discontinuity at* $x = a$ *for* $a \neq 0$*. So discussion below will give you a method to tackle atleast some type of jump discontinuities.*

So consider

$$\varphi_a(x) \;=\; (x-a)I_{(-\infty,\,a]}(x).$$

Note φ_a is continuous and hence using Theorem 0.1, we get

$$\begin{aligned} E[(X-a)I_{\{X\le a\}}] &= \int_{-\infty}^{\infty}(x-a)I_{(-\infty,\,a]}(x)f(x)dx \\ &= \int_{-\infty}^{a}(x-a)f(x)dx \\ &= \int_{-\infty}^{a}xf(x)dx - aP\big(X\le a\big). \end{aligned}$$

Now

$$\begin{aligned} E[(X-a)I_{\{X\le a\}}] &= E[XI_{\{X\le a\}}] - aE[I_{\{X\le a\}}] \\ &= E[XI_{\{X\le a\}}] - aP\big(X\le a\big). \end{aligned}$$

Now equating the above two, we get

$$E[XI_{\{X\le a\}}] = \int_{-\infty}^{a} xf(x)dx.$$

Exercise Let φ be a continuous function and $a\in\mathbb{R}$ and X be a random variable with pdf f such that $E[\varphi(X)]$ is finite. Show that

$$E[\varphi(X)I_{\{X\le a\}}] = \int_{-\infty}^{a}\varphi(x)f(x)dx.$$

Along similar line to Theorem 0.1, we have the following theorem.

Theorem 0.2 *Let X and Y be continuous random variables with joint pdf f and $\varphi:\mathbb{R}^2\to\mathbb{R}$ be a continuous function. Then*

$$E[\varphi\circ(X,Y)] \;=\; \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\varphi(x,y)f(x,y)dxdy.$$

Here again, proof is beyond the scope of the course but special cases like $X+Y, XY, X^2$ etc . I will do the case $\varphi(x,y)=xy$ as an example.

First note that the pdf g of $Z=XY$ is given by

$$g(z) \;=\; \int_{-\infty}^{\infty}\frac{1}{|x|}f(x,\frac{z}{x})dx, z\in\mathbb{R}.$$

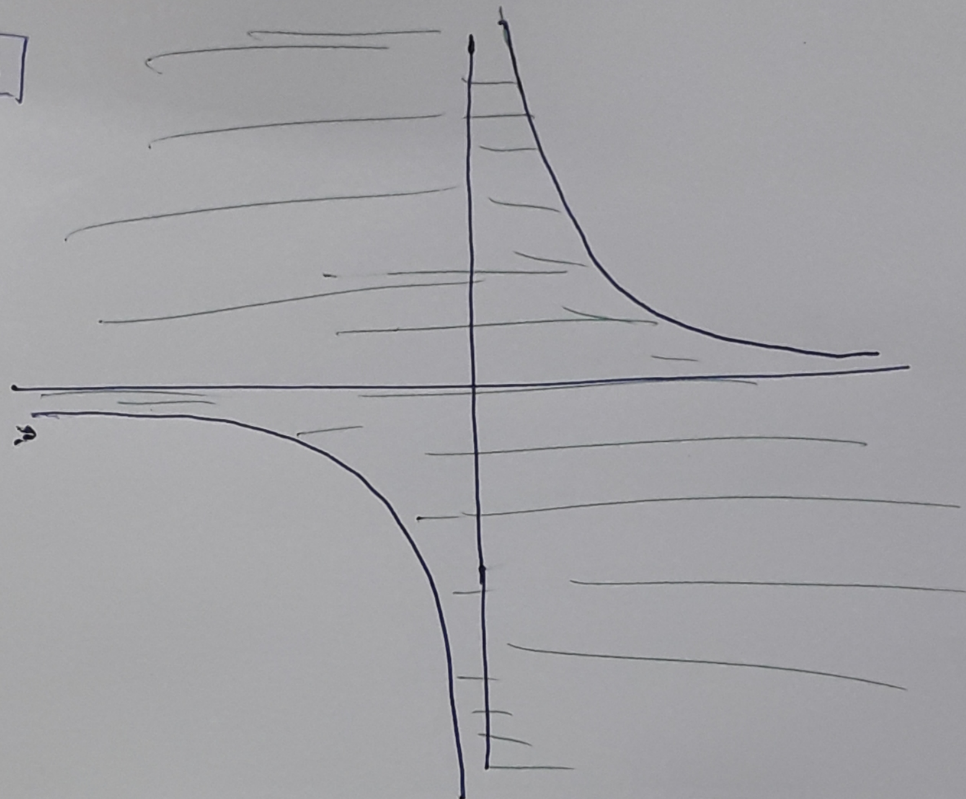
(If you have not yet done this, immediately attend the above.)
Note that

$$P(XY \leq z) = P((X,Y) \in A_z)$$

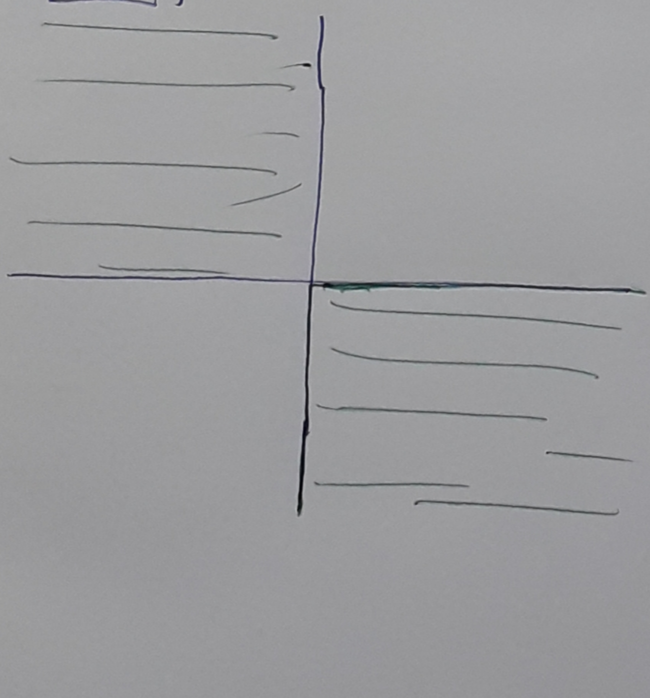
where A_z is as below. Observe the way, the region changing from positive z to negative.

$$A_z = \{ (x,y) \mid xy \leq z \}$$

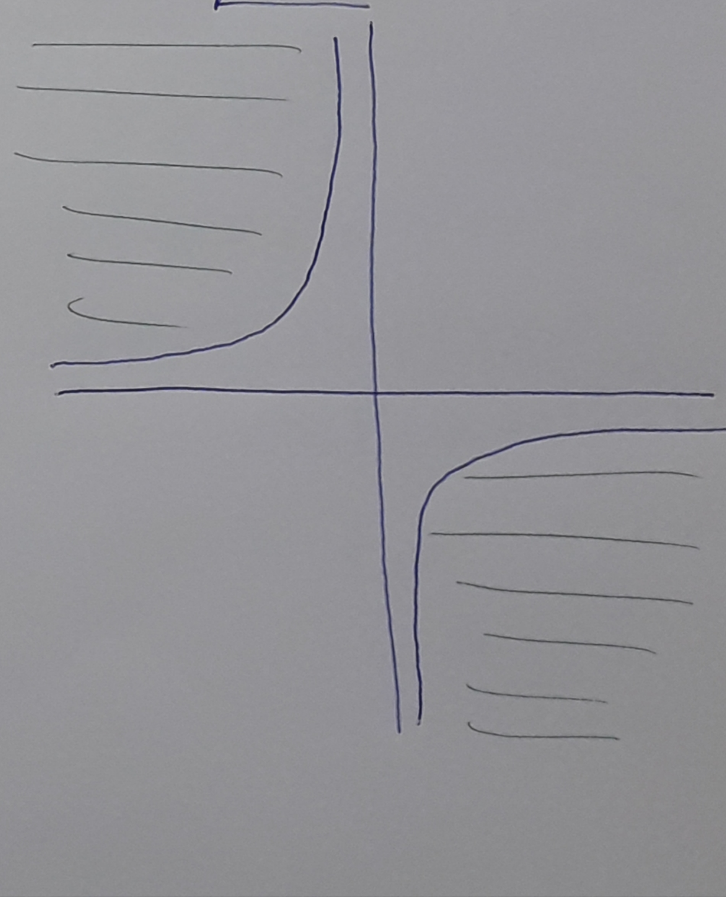
$\boxed{z=1}$



$\boxed{z=0}$



$\boxed{z=-1}$



Hence

$$\begin{aligned}
E[XY] &= \int_{-\infty}^{\infty} zg(z)dz \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{z}{|x|} f(x, \frac{z}{x})dxdz \\
\text{(change order of integrtaion)} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{z}{|x|} f(x, \frac{z}{x})dzdx \\
\text{(put } z = xy) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{xy}{|x|} f(x,y)|x|dydx \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dydx.
\end{aligned}$$

Theorem 0.3 *Let X and Y be independent random variables such that EX, EY exists. Then $E[XY]$ exists and is given by*

$$E[XY] = EXEY.$$

Proof: (Special cases) Using the above one can see that when X and Y are independent and with a pdf f, then

$$E[XY] = EXEY$$

(exercise). When X, Y discrete with joint pmf f, then again it is easy to prove (exercise). □

Variance and other Higer order moments: In this subsection, we address the question of recontruction of the distribution of a random variable. To this end, we introduce objects namely moments of a random variable and later see whether we can determine the distribution function using them.

Definition 7.5.(Higher Order Moments) Let X be a random variable. Then EX^n is called the nth moment of X and $E(X - EX)^n$ is called the nth central moment of X. The second central moment is called the variance.

Theorem 0.4 *(Markov inequality) Let X be a non negative random variable with finite nth moment. Then we have for each $\Delta > 0$*

$$P\big(X \geq \Delta\big) \leq \frac{EX^n}{\Delta^n}.$$

Proof. Set

$$Y(\omega) \;=\; \begin{cases} 0 & \text{if} \quad X(\omega) < \Delta \\ \Delta^n & \text{if} \quad X(\omega) \geq \Delta \,. \end{cases}$$

Then Y is a non negative simple random arable such that $Y \leq X^n$. Hence from Theorem 0.3, Lecturenotes18-19, we have $EY \leq EX^n$. Therefore

$$\Delta^n \, P\big(X \geq \Delta\big) \;=\; EY \leq \; EX^n \,.$$

This completes the proof. □

Example 0.3 *In the Markov inequality, non negativity is important. For example, take $X = 1$ or -1 with probability $\frac{1}{2}$. Then Markov inequality doesn't hold. For example take $\Delta = \frac{1}{2}$ and $n = 1$ and note that $EX = 0$.*

As a corollary we have the Chebyshev's inequality.

Chebyshev's inequality. Let X be a random variable with finite mean μ and finite variance σ^2. Then for each $\Delta > 0$,

$$P\big(|X - \mu| \geq \Delta\big) \;\leq\; \frac{\sigma^2}{\Delta^2}$$

The proof of Chebyshev's inequality follows by replacing X by $|X - \mu|$ in the Markov inequality.

Example 0.4 *(1) Let X be Bernoulli (p). Then we have see that $EX = p$ and hence for a Bernoulli random variable, knowing first moment itself will uniquely identify the distribution. Also note that other moments are $EX^n = p, n \geq 1$. Observe the pattern of the moments, $\{p, p, \cdots\}$.*

(2) Let X be Binomial (n, p). Then $EX = np$. This doesn't gives uniquely the distribution. So let us compute the variance. To do this, we use the following. Let $X_1, X_2, \cdots, X_n$ be n-independent Bernoulli(p) random variables. Then we know that $X_1 + \cdots + X_n$ is Binomial (n, p). Hence take $X = X_1 + \cdots + X_n$. Now

$$\begin{aligned} E[X - EX]^2 &= \sum_{k=1}^{n} E[X_k - p]^2 \\ &= \sum_{k=1}^{n} p(1-p) \\ &= np(1-p). \end{aligned}$$

Now given EX and $E[X-EX]^2$, we can solve

$$EX = np,\ E[X-EX]^2 = np(1-p)$$

to find the parameters n and p.

Also find few more moments (exercise).

(3) *Let $X \sim N(0,1)$. Then $EX = 0$ and $EX^2 = 1$ and also Variance* $\mathrm{Var}(X) = 1$. *(Exercise. Hint: To find the variance of $Y \sim N(0,1)$, use integration by parts to $\int_{-\infty}^{\infty} 1.e^{-\frac{x^2}{2}} dx$.)*

The above examples only tell that if we know apriori that given random variable is of certain type like Binomial, Bernoulli, Poisson but we don't know the parameters, one can use their moments to determine their parameters and hence their distribution.

There is an interesting problem, given a sequence of numbers $\{a_1, a_2, \cdots\}$, can one find a distribution/ distribution function whose moments are given by $\{a_n\}$? Also is it unique. This is called the moment problem.

For example, we have seen that the sequence $\{p, p, \cdots\}$ corresponds to the Bernoulli (p) distribution but not sure right now whether it is the only distribution with moments given by the above sequence.

If a distribution is uniquely determined by its moments are called moment determinant distributions and others are called moment indeterminant.

In fact, Bernoulli, Binomial, Poisson, Normal etc are moment determinant but log normal distribution is moment indeterminant.

0.1 Stirling's formula

In this section, we prove the following asymptotic formula which gives an approximation for $n!$.

First we introduce the Lapalce method. Method enables one to approximate integrals of the form $\int e^{nf(x)}dx$ in terms of Gaussian integrals when n is very large if f is a strictly concave function with unique maximum at x_0 with $f''(x_0) < 0$. A heuristic description of the method is as follows.

$$\begin{aligned} \int_a^\infty e^{nf(x)}dx &= e^{nf(x_0)} \int_a^\infty e^{n(f(x)-f(x_0))}dx \\ &\approx e^{nf(x_0)} \int_a^\infty e^{\frac{n}{2}f''(x_0)(x-x_0)^2}dx \end{aligned}$$

i.e. for n very large,

$$\begin{aligned}\int_a^\infty e^{nf(x)}dx &\approx e^{nf(x_0)}\int_a^\infty e^{-\frac{n}{2}|f''(x_0)|(x-x_0)^2}dx \\ &\approx \frac{\sqrt{2\pi}}{\sqrt{n|f''(x_0)|}}e^{nf(x_0)}\,\frac{\sqrt{n|f''(x_0)|}}{\sqrt{2\pi}}\int_{-\infty}^\infty e^{-\frac{n}{2}|f''(x_0)|(x-x_0)^2}dx\end{aligned}$$

In the above note that $\frac{\sqrt{n|f''(x_0)|}}{\sqrt{2\pi}}e^{-\frac{n}{2}|f''(x_0)|(x-x_0)^2}$ is the pdf of Normal random variable with mean x_0 and variance $\frac{1}{n|f''(x_0)|}$. Hence

$$\int_a^\infty e^{nf(x)}dx \approx \frac{\sqrt{2\pi}}{\sqrt{n|f''(x_0)|}}e^{nf(x_0)} \text{ for large enough } n.$$

Now let us give a formal statement.

Lemma 0.1 *Let $f:(a,\ \infty)\to\mathbb{R},\ a\in\mathbb{R}$ be twice continuously differentiable function with unique maximum at x_0, satifies $f''(x_0)<0$ and $\int_a^\infty e^{f(x)}dx<\infty$. Then*

$$\lim_{n\to\infty}\frac{\sqrt{n|f''(x_0)|}}{\sqrt{2\pi}\,e^{nf(x_0)}}\int_a^\infty e^{nf(x)}dx = 1.$$

Proof: (Reading exercise)
Claim 1

$$\lim_{n\to\infty}\frac{\sqrt{n|f''(x_0)|}}{\sqrt{2\pi}\,e^{nf(x_0)}}\int_a^\infty e^{nf(x)}dx \geq 1.$$

<u>Proof</u> : For $\varepsilon>0$, there exists $\delta>0$ such that

$$|f''(x)-f''(x_0)|<\varepsilon \text{ whenever } |x-x_0|<\delta.$$

This implies

$$f''(x)>f''(x_0)-\varepsilon \text{ whenever } |x-x_0|<\delta.$$

Now using Taylor series expansion, we have

$$f(x)=f(x_0)+\frac{(x-x_0)}{2}f''(c)$$

where c lies between x and x_0. Combining the above two, we get for $x\in(x_0-\delta,\ x_0+\delta)$,

$$f(x)>f(x_0)+\frac{(x-x_0)}{2}(f''(x_0)-\varepsilon). \tag{0.1}$$

Now

$$\begin{aligned}
\int_a^\infty e^{nf(x)}dx &\geq \int_{x_0-\delta}^{x_0+\delta} e^{nf(x)}dx \\
&\geq e^{nf(x_0)} \int_{x_0-\delta}^{x_0+\delta} e^{\frac{n(f''(x_0)-\varepsilon)}{2}(x-x_0)^2}dx \\
\big(\text{use } y=(x-x_0)\sqrt{n(|f''(x_0)|+\varepsilon)}\big) \quad &= e^{nf(x_0)} \frac{1}{\sqrt{n(|f''(x_0)|+\varepsilon)}} \int_{-\delta\sqrt{n(|f''(x_0)|+\varepsilon)}}^{\delta\sqrt{n(|f''(x_0)|+\varepsilon)}} e^{-\frac{y^2}{2}}dy
\end{aligned}$$

Hence

$$\frac{\sqrt{n|f''(x_0)|}}{\sqrt{2\pi}\, e^{nf(x_0)}} \int_a^\infty e^{nf(x)}dx \quad \geq \quad \frac{\sqrt{|f''(x_0)}}{\sqrt{(|f''(x_0)|+\varepsilon)}} \frac{1}{\sqrt{2\pi}} \int_{-\delta\sqrt{n(|f''(x_0)|+\varepsilon)}}^{\delta\sqrt{n(|f''(x_0)|+\varepsilon)}} e^{-\frac{y^2}{2}}dy$$

Since

$$\lim_{n\to\infty} \frac{1}{\sqrt{2\pi}} \int_{-\delta\sqrt{n(|f''(x_0)|+\varepsilon)}}^{\delta\sqrt{n(|f''(x_0)|+\varepsilon)}} e^{-\frac{y^2}{2}}dy \;=\; \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}dy \;=1$$

we get

$$\lim_{n\to\infty} \frac{\sqrt{n|f''(x_0)|}}{\sqrt{2\pi}\, e^{nf(x_0)}} \int_a^\infty e^{nf(x)}dx \geq \frac{\sqrt{|f''(x_0)}}{\sqrt{(|f''(x_0)|+\varepsilon)}},\ \varepsilon>0.$$

Since $\varepsilon > 0$ is chosen arbitrarily, we get

$$\lim_{n\to\infty} \frac{\sqrt{n|f''(x_0)|}}{\sqrt{2\pi}\, e^{nf(x_0)}} \int_a^\infty e^{nf(x)}dx \geq 1.$$

Claim 2

$$\lim_{n\to\infty} \frac{\sqrt{n|f''(x_0)|}}{\sqrt{2\pi}\, e^{nf(x_0)}} \int_a^\infty e^{nf(x)}dx \leq 1.$$

Proof (similar argument- so reading exercise) Since $f''(x_0) < 0$, for $\varepsilon > 0$ small enough, $f''(x_0) + \varepsilon < 0$. for such $\varepsilon > 0$, there exists $\delta > 0$ such that

$$f''(x) < f''(x_0) + \varepsilon,\ \text{whenever } |x - x_0| < \delta.$$

Hence by using Taylor series exapansion around x_0, we get

$$f(x) \leq f(x_0) + \frac{1}{2}(x-x_0)^2(f''(x_0)+\varepsilon),\ x \in (x_0-\delta,\ x_0+\delta). \qquad (0.2)$$

Now

$$\int_a^\infty e^{nf(x)}dx = \int_a^{x_0-\delta} e^{nf(x)}dx + \int_{x_0-\delta}^{x_0+\delta} e^{nf(x)}dx + \int_{x_0+\delta}^\infty e^{nf(x)}dx. \quad (0.3)$$

Since x_0 is a strict maximum of f, there exists $\eta > 0$ such that

$$f(x) \le f(x_0) - \eta, \text{ for } |x - x_0| \ge \delta. \quad (0.4)$$

Using (0.4),

$$\begin{aligned} \int_a^{x_0-\delta} e^{nf(x)}dx + \int_{x_0+\delta}^\infty e^{nf(x)}dx &\le \int_a^\infty e^{f(x)} e^{(n-1)(f(x_0)-\eta)}dx \\ &= Me^{(n-1)(f(x_0)-\eta)}, \end{aligned} \quad (0.5)$$

where

$$M = \int_a^\infty e^{f(x)}dx.$$

Now as in the proof of Claim 1, we get

$$\int_{x_0-\delta}^{x_0+\delta} e^{nf(x)}dx \le e^{nf(x_0)} \frac{\sqrt{2\pi}}{\sqrt{n(|f''(x_0) - \varepsilon)}}. \quad (0.6)$$

Hence combining (0.3), (0.5) and (0.6) we get

$$\int_a^\infty e^{nf(x)}dx \le e^{nf(x_0)} \Big[Me^{-f(x_0)} e^{-(n-1)\eta} + \frac{\sqrt{2\pi}}{\sqrt{n(|f''(x_0)| - \varepsilon)}} \Big].$$

Hence

$$\frac{\sqrt{n|f''(x_0)|}}{\sqrt{2\pi}\, e^{nf(x_0)}} \int_a^\infty e^{nf(x)}dx \le \frac{Me^{-f(x_0)}\sqrt{|f''(x_0)|}}{\sqrt{2\pi}} \sqrt{n} e^{-(n-1)\eta} + \frac{|f''(x_0)|}{\sqrt{|f''(x_0| - \varepsilon}}$$

Therefore

$$\lim_{n\to\infty} \frac{\sqrt{n|f''(x_0)|}}{\sqrt{2\pi}\, e^{nf(x_0)}} \int_a^\infty e^{nf(x)}dx \le \frac{|f''(x_0)|}{\sqrt{|f''(x_0| - \varepsilon}}.$$

Since $\varepsilon > 0$ is arbitrarily small, we get

$$\lim_{n\to\infty} \frac{\sqrt{n|f''(x_0)|}}{\sqrt{2\pi}\, e^{nf(x_0)}} \int_a^\infty e^{nf(x)}dx \le 1.$$

Combining Claims 1 and 2, we get the result. $\square$

Lemma 0.2 *(Stirling's formula)*

$$\lim_{n\to\infty} \frac{e^n\, n!}{n^n \sqrt{n}} \;=\; \sqrt{2\pi}.$$

Proof:

Consider

$$\begin{aligned}\int_0^\infty x^n e^{-x} dx &= -x^n e^{-x}\Big|_0^\infty + n\int_0^\infty x^{n-1} e^x dx \\ \text{(repeat the integration by parts)} &= n! \int_0^\infty e^{-x} dx = n!.\end{aligned}$$

Hence

$$\begin{aligned} n! &= \int_0^\infty x^n e^{-x} dx \\ &= \int_0^\infty e^{n\ln x - x} dx \\ \text{(use substitution } x = ny) &= ne^{n\ln n} \int_0^\infty e^{n\ln y - ny} dy \\ &= ne^{n\ln n} \int_0^\infty e^{n(\ln x - x)} dx. \qquad (*)\end{aligned}$$

We use Lemma 0.1 to $f(x) = \ln x - x$. Note that $f(x) = \ln x - x, x > 0$ has a unique maximum at $x = 1$ and $\int e^{\ln x - x} dx < \infty$. $f''(1) = -1, f(1) = -1$ So using Lemma 0.1, we get

$$\lim_{n\to\infty} \frac{\sqrt{n}}{\sqrt{2\pi} e^{-n}} \int_0^\infty e^{n(\ln x - x)} dx = 1. \qquad (0.7)$$

Therefore

$$\lim_{n\to\infty} \frac{e^n\, n!}{n^n \sqrt{n}} \;=\; \sqrt{2\pi} \lim_{n\to\infty} \frac{\sqrt{n} e^n}{\sqrt{2\pi}} \int_0^\infty e^{n(\ln x - x)} dx \;=\; \sqrt{2\pi}.$$

□

Chapter 7 : Moment generating function and Characteristic function

Key words: Characteristic function, inverse theorem, uniqueness theorem, continuity theorem.

In this chapter, we introduce the notion of moment generating function (in short mgf) and characteristic function of a random variable and study its properties. Both moment generating function and Characteristic function can be used to identify distribution functions uniquely unlike moments. In fact, a way to understand whether a distribution is moment determinant or not is by using either moment generating function or characteristic functions. It is interesting to note that mgf is closely related the Laplace transform and characteristic function is its counter part Fourier transform.

0.1.1 Moment generating function

In this subsection we study moment generating functions and its properties.

Definition 0.1 *Given a random variable on a probability space* $(\Omega, \mathcal{F}, P)$*, its moment generating function denoted by* M_X *is defined as*

$$M_X(t) = E[e^{tX}], t \in I,$$

where I *is an interval on which the rhs expectation exists.*

In fact for a non negative random variable X, I always contains $(-\infty, 0]$. There are non negative random variables X such that $M_X(t)$ doesn't exists for $t > 0$. Will see examples soon. Analogous comment holds for negative random variable. Moment generating functions becomes special when I contains an interval containing 0.

Example 0.5 *Let* $X \sim$ *Bernoulli* (p)*. Then*

$$M_X(t) \;=\; (1-p) + pe^t, \; t \in \mathbb{R}.$$

Now we will state and indicate the proofs of various properties of moment generating functions.

Theorem 0.5 *Let* X *be a random with mgf* $M_X(t), t \in I$ *and* $Y = aX + b, a \neq 0$*. Then the mgf of* Y *is given by*

$$M_Y(t) \;=\; e^{bt} M_X(t), t \in J,$$

where $J = \{t \in \mathbb{R} | at \in I\} := a^{-1}I$*.*

Proof: For $t \in a^{-1}I$,

$$\begin{aligned} E[e^{tY}] &= E[e^{bt}e^{atX}] \\ &= e^{bt}M_X(at). \end{aligned}$$

Note that for $t \in a^{-1}I,\ at \in I$ and hence $M_X(at)$ exists. $\square$

Remark 0.1 *In Theorem 0.5, the result holds when $Y \sim aX + b$, i.e. if $Y \sim aX + b, a \neq 0$, then*

$$M_Y(t) = e^{bt}M_X(at), t \in J.$$

Following is an illustration of the use of the above theorem.

Example 0.6 *Let $X \sim N(\mu, \sigma^2)$. Then $X = \mu + \sigma Y$, $Y \sim N(0, 1)$. Now*

$$\begin{aligned} M_Y(t) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{ty}e^{-\frac{1}{2}y^2}dy \\ &= e^{\frac{t^2}{2}}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(y-t)^2}dy \\ &= e^{\frac{t^2}{2}}, t \in \mathbb{R}. \end{aligned}$$

i.e. mgf of standard normal is $e^{\frac{t^2}{2}}, t \in \mathbb{R}$. Now using the above theorem

$$M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}, t \in \mathbb{R}^2.$$

Theorem 0.6 *Let X be a random variable such that $M_X(t)$ exists in an interval I which contains $[-h, h]$ for some $h > 0$. Then X has moments of all orders and the mgf M_X has all derivatives on $(-h, h)$ and the following holds.*

$$EX^k = M_X^{(k)}(0),\ k = 0, 1, \cdots.$$

Here $M^{(k)}(t)$ denote the kth derivative of $M_X(t)$ at t. Proof follows from

$$\frac{d^kE[e^{tX}]}{dt^k} = E\Big[\frac{d^ke^{tX}}{dt^k}\Big], t \in (-h, h).$$

So proof is all about justifying differentation under the 'integral' sign.

The above theorem points to the following. Existence of mgf in a neighourhood of 0, makes the random variable very nice. This implies an unpleasent property of mgfs, i.e. they won't exists arround zero unless the random variable is nice". We will see some examples to illustrate this point.

Example 0.7 *(Cauchy distribution) Let X be a Cauchy random variable, i.e. X is a continuous random variable with pdf*

$$f(x) = \frac{1}{\pi(1+x^2)}, \ x \in \mathbb{R}.$$

We know that X doesn't have finite mean (exercise). Hence Theorem 0.6 points to the non existence of mgf. In fact, we show that mgf of X exists only at $t = 0$. For $t > 0$, consider

$$\begin{aligned} E[e^{tX}] &= \frac{1}{\pi}\int_{-\infty}^{\infty} e^{tx}\frac{1}{1+x^2}dx \\ &\geq \frac{1}{\pi}\int_{0}^{\infty} e^{tx}\frac{1}{1+x^2}dx \\ &\geq \frac{t}{\pi}\int_{0}^{\infty} \frac{x}{1+x^2}dx \\ &= \frac{t}{2\pi}\int_{1}^{\infty} \frac{1}{y}dy. \end{aligned}$$

The RHS integral diverges to ∞. Hence by comparison of integrals, it follows that $E[e^{tX}]$ diverges to ∞. i.e. $M_X(t)$ doesn't exists for $t > 0$. Using a simliar argument one can show that $M_X(t)$ doesn't exists for $t < 0$ (exercise). So $M_X(t)$ exists only at 0.

Example 0.8 *(Log normal distribution) Let $X \sim e^Y$, $Y \sim N(\mu, \sigma^2)$. Then X is said to be 'log normally' distributed.*

Take $\mu = 0, \sigma = 1$.

$$EX^n = E[e^{nY}] = M_Y(n) = e^{\frac{n^2}{2}}, n \geq 1.$$

Now since $X \geq 0$, clearly $M_X(t)$ exists for $t \geq 0$. Now for $t > 0$,

$$\begin{aligned} E[e^{tX}] &= E[e^{te^Y}] \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{te^y}e^{-\frac{y^2}{2}}dy \\ &\geq \frac{1}{\sqrt{2\pi}}\int_{K}^{\infty} e^{t}dy, \end{aligned}$$

for some $K > 0$ large enough. In the first inequality we used the fact that there exists a $K > 0$

$$te^y - \frac{y^2}{2} \geq t \text{ for all } y \geq K.$$

[1] *Now since the last integral diverges to ∞, it follows that $E[e^{tX}]$ diverges to ∞. Hence $M_X(t)$ doesn't exists for $t > 0$, though EX^n exists for all $n \geq 1$.*

Theorem 0.7 *Let $X_1, X_2, \cdots, X_n$ be indedependent random variables with mgfs $M_{X_i}(t)$ exists on a common interval I, then the mgf of the sum $S_n = X_1 + \cdots + X_n$ exists on I and is given by*

$$M_{S_n}(t) = \Pi_{k=1}^n M_{X_k}(t), t \in I.$$

Proof: We use the following results.

- If X and Y are independent, then $E[XY] = EXEY$.
- If X and Y are independent so are $f \circ X$ and $g \circ Y$, where f, g are Borel functions.

Now

$$\begin{aligned} E[e^{tS_n}] &= E[e^{tX_1} \cdots e^{tX_n}] \\ &= \Pi_{i=1}^n E[e^{tX_i}] \\ &= \Pi_{i=1}^n M_{X_i}(t), \ t \in I. \end{aligned}$$

This completes the proof. □

Example 0.9 *Let $X \sim$ Binomial (n, p). Note that $X = X_1 + \cdots + X_n$, where X_i's are i.i.d. Bernoulli (p). Hence*

$$M_X(t) = (M_{X_1}(t))^n = (1 - p + pe^t)^n, t \in \mathbb{R}.$$

Theorem 0.8 *(moment determinant distributions) Let X be a random variable such that the moments $\mu_k = EX^k, k \geq 1$ exists and satisfies*

$$\lim_{k \to \infty} \frac{1}{2k} \mu_{2k}^{\frac{1}{2k}} = 0.$$

Then if Y is another random variable with $EY^k = \mu_k$ for all $k \geq 1$, then X and Y has same distribution.

---

[1]For $0 < t < 1$,

$$\begin{aligned} te^y - \frac{y^2}{2} &\geq t(1 + y + y^2/2 + y^3/6) - y^2/2 \\ &= t + ty + y^2/2(t - 1 + t/3y) \\ &\geq t \text{ for } y > \frac{3(1-t)}{t}. \end{aligned}$$

For $t > 1$, $te^y - \frac{y^2}{2} \geq t$ for all y.

The proof follows from the Riesz criterion (sufficient condition) for moment determinancy given by

$$\liminf_{k\to\infty}\left(\frac{\mu_{2k}}{(2k)!}\right)^{\frac{1}{2k}} < \infty.$$

and the Stirling's approximation given by

$$\lim_{n\to\infty}\frac{n!}{\sqrt{2\pi}n^{n+\frac{1}{2}}e^{-n}} \;=\; 1.$$

Remark 0.2 *Theorem 0.8 gives a partial converse for Theorem 0.6. i.e., " if $\mu_k = EX^k$ exists for all k, then $M_X(t)$ is uniquely determined by μ_k's" is not true in general. But with an extra condition that μ_k doesn't grow rapidly (for example if $\lim_{k\to\infty}\frac{1}{2k}\mu_{2k}^{\frac{1}{2k}} = 0$, then mgf is uniquely determined. This gives a partial answer to the question, when given all moments determine a distribution uniquely, because mgfs determinds distributions uniquely, a fact we will not state or prove in this course.*

Now we will see some more examples.

Example 0.10 *Let $X \sim$ Binomial (n,p). Then set*

$$P\big(X = k\big) \;=\; \binom{n}{k}p^k(1-p)^{n-k} := p_k,\;\; k = 0,1,\cdots,n.$$

$$\mu_m \;=\; EX^m \;\le n^{2m}.$$

Therefore

$$\lim_{m\to\infty}\frac{1}{2m}(\mu_{2m})^{\frac{1}{2m}} \;=\; 0.$$

i.e. Binomial (n,p) is moment determinant.

Example 0.11 *Let $X \sim N(0,1)$. Then*

$$\mu_{2m} \;=\; 1\cdot 3\cdot\cdots\cdot(2m-1) \;=\; \frac{(2m)!}{2^m m!}.$$

Hence

$$\begin{aligned}
\lim_{m\to\infty}\frac{1}{2m}\mu_{2m}^{\frac{1}{2m}} &= \frac{1}{\sqrt{2}}\lim_{m\to\infty}\frac{1}{2m}\left[\frac{2m!}{m!}\right]^{\frac{1}{2m}}\\
(\text{Stirling's approx.}) &= \frac{1}{\sqrt{2}}\lim_{m\to\infty}\frac{1}{2m}\left[\frac{\sqrt{2\pi}(2m)^{2m+\frac{1}{2}}e^{-2m}}{\sqrt{2\pi}m^{m+\frac{1}{2}}e^{-m}}\right]^{\frac{1}{2m}}\\
&= 0.
\end{aligned}$$

Example 0.12 *Let X be such that $X \sim e^Y$, $Y \sim N(0,1)$. The X has a pdf given by*

$$f(x) \;=\; \frac{1}{\sqrt{2\pi}x} e^{-\frac{1}{2}(\log x)^2},\; x > 0,\; = 0,\; x \leq 0.$$

Now set

$$g(x) \;=\; f(x)\big(1 + \sin(2\pi \log x)\big),\; x \geq 0,\; = 0,\; x \leq 0.$$

Now consider

$$\begin{aligned}
\int_0^\infty x^n f(x) \sin(2\pi \log x) dx \;&=\; \int_{-\infty}^\infty e^{ny} f(e^y) \sin(2\pi y) e^y dy \\
&=\; \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{ny} \sin(2\pi y) e^{-\frac{1}{2}y^2} dy \\
&=\; e^{\frac{1}{2}n^2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty \sin 2\pi y e^{-\frac{1}{2}(y-n)^2} dy \\
&=\; e^{\frac{1}{2}n^2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty \sin 2\pi(x+n)\, e^{-\frac{1}{2}x^2} dx \\
&=\; e^{\frac{1}{2}n^2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty \sin 2\pi x\, e^{-\frac{1}{2}x^2} dx \\
&=\; 0,
\end{aligned}$$

the first equality follows using change of variable formula with $y = \log x, x > 0$ and the last equality follows from the fact that integrand is an odd function which is absolutely integrable. Hence

$$\int_0^\infty x^n g(x) dx \;=\; \int_0^\infty x^n f(x) dx,\; n \geq 0.$$

In the above, $n = 0$ implies that g is a pdf and let Z be a random variable with pdf g, the from the above $EX^n = EZ^n$ for all $n \geq 1$. i.e. X is not moment determinant.

0.1.2 Characteristic functions

Definition 0.2 *(Characteristic functions) The characteristic function of a random variable X is defined as*

$$\Phi_X(t) \;=\; Ee^{itX}, t \in \mathbb{R}.$$

(where $Ee^{itX} \;=\; E\cos tX + iE\sin tX$).

A digression: Before going further, I will give some very brief working knowledge for complex valued functions defined on $\mathbb{R}$.

- Given a function $f = f_1 + if_2 : \mathbb{R} \to \mathbb{C}$, f_1, f_2 are the real and imaginary parts of f.

- Given two functions $f, g : \mathbb{R} \to \mathbb{C}$ with $f = f_1 + if_2, g = g_1 + ig_2$, we define

$$\begin{aligned}
(f+g)(x) &:= f_1(x) + g_1(x) + i(f_2(x) + g_2(x)), x \in \mathbb{R}, \\
(fg)(x) &= f(x)g(x) \\
&= f_1(x)g_1(x) - f_2(x)g_2(x) + i(f_1(x)g_2(x) + f_2(x)g_1(x)), x \in \mathbb{R} \\
\left(\frac{1}{f}\right)(x) &= \frac{1}{f_1(x) + if_2(x)}, x \in \mathbb{R}.
\end{aligned}$$

- Given a finction $g : \mathbb{R} \to \mathbb{C}$, we say that g is continuous at x if both g_1 and g_2 are continuous at x. This equivalent to the following. "For each $\varepsilon > 0$, there exists $\delta > 0$ such that

$$|x - y| < \delta \Rightarrow \ |g(x) - g(y)| < \varepsilon.$$

- A function $g : \mathbb{R} \to \mathbb{C}$ is uniformly continuous if g_1 and g_2 are uniformly continuous. This is equivalent to "for each $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$\text{for all } x, y \in \mathbb{R}, \text{ with } |x - y| < \delta \ \Rightarrow \ |g(x) - g(y)| < \varepsilon.$$

- $g : \mathbb{R} \to \mathbb{C}$ is said be differentable at x, if g_1 and g_2 are differentiable at x and in this case $g'(x) \ = \ g_1'(x) + ig_2'(x)$.

- $g : \mathbb{R} \to \mathbb{C}$ is (Riemann) integrable on $[a,\ b]$ if g_1 and g_2 are integrable on $[a,\ b]$ and in this case

$$\int_a^b g(x)dx \ = \ \int_a^b g_1(x)dx + i\int_a^b g_2(x)dx.$$

  These are special class of line integrals.

- Let $G : \mathbb{R} \to \mathbb{C}$ be the primitive of $g : \mathbb{R} \to \mathbb{C}$, i.e., $(G'(x) = g(x), x \in \mathbb{R})$, then for any $a < b$,

$$\int_a^b g(x)dx \ = \ G(b) - G(a).$$

  (This is called the fundamental theorem for line integrals).

Example 0.13 *Let* $X \sim \text{Bernoulli}(p)$. *Then*

$$\phi_X(t) \;=\; (1-p) + pe^{it}\,.$$

Example 0.14 *Let* $X \sim \text{exponential}(\lambda),\ \lambda > 0$. *Then*

$$\begin{aligned}
\Phi_X(t) \;=\; Ee^{itX} &= \lambda \int_0^\infty e^{itx} e^{-\lambda x} dx \\
&= \int_0^\infty e^{(it-\lambda)x} dx \\
&= \tfrac{\lambda}{it-\lambda} \left[e^{(it-\lambda)x} \right]_0^\infty \\
&= \tfrac{\lambda}{it-\lambda},\ t \in \mathbb{R}.
\end{aligned}$$

In the third equality, we used fundamental theorem for line integrals and in the fourth equality, we used

$$\lim_{x\to\infty} e^{(it-\lambda)x} \;=\; 0.$$

Theorem 0.9 *For any random variable* X*, its characteristic function* $\phi_X(\cdot)$ *is uniformly continuous on* $\mathbb{R}$ *and satisfies*
(i) $\Phi_X(0) \;=\; 1$
(ii) $|\Phi_X(t)| \le 1$
(iii) $\Phi_X(-t) \;=\; \overline{\Phi_X(t)}$, *where for* z *a complex number,* $\overline{z}$ *denote its conjugate.*

Proof:
We prove (iii), (i) and (ii) are exercises.

$$\begin{aligned}
\Phi_X(-t) \;=\; Ee^{-itX} &= E\cos tX - iE\sin tX \\
&= \overline{E\cos tX + iE\sin tX} \\
&= \overline{\Phi_X(t)}\,.
\end{aligned}$$

Now we show that Φ_X is uniformly continuous. Consider

$$\begin{aligned}
|\Phi_X(t+h) - \Phi_X(t)| &= |E(e^{i(t+h)X} - e^{itX})|, \\
&\le E|e^{ihX} - 1| \\
&= E\sqrt{2(1-\cos(hX))} \\
&= 2E|\sin(\tfrac{hX}{2})|
\end{aligned}$$

Note that

$$\lim_{h\to 0} \left| \sin\left(\frac{hX(\omega)}{2} \right) \right| = 0, \text{ and } \left| \sin(\frac{hX}{2}) \right| \le 1.$$

Hence, using Dominated Convergence theorem, $\Phi_X(t+h) \to \Phi_X(t)$ uniformly in t as $h \to 0$. This imply that Φ_X is uniformly continuous.