Chapter 3. Continuous Distributions

Objectives

- Basic Concepts & Expectations
- Uniform, Exponential, Γ, χ^2, and Normal Distributions
- Introduction to the Maximum Likelihood Estimation
- The Central Limit Theorem

Continuous Distributions

Definition 1. *$X \sim$ a random variable of continuous type*
If for each event $A \subset S$, $P(A)$ can be written as

$$P(A) = P(X \in A) = \int_A f(x)\,dx$$

for some function f with $f(\cdot) > 0$ on S, X is a random variable of continuous type and f is called the pdf of X.

Properties 1.

1. $f(x) > 0, \quad \forall x \in S$
2. $\displaystyle\int_S f(x)dx = 1$
3. $\displaystyle P(a < X < b) = \int_a^b f(x)\,dx, \quad for\,(a,b) \subset S$

Ex 1. Suppose $X \sim uniform\,[0,1)$, i.e., pdf $f(x) = 1, \quad 0 \le x < 1$. Find (a) $P(0.1 < X < 0.4)$, (b) $P(X = 0.5)$.

Answer:

$$P(0.1 < X < 0.4) = \int_{0.1}^{0.4} f(x)dx = x\Big|_{0.1}^{0.4} = 0.3$$
$$P(X = 0.5) = \int_{0.5}^{0.5} f(x)dx = 0$$

Ex 2. Suppose $X \sim exponential\,(20)$, i.e., pdf $f(x) = \frac{1}{20}e^{-x/20}, \quad 0 \le x < \infty$. Find $P(X > 20)$.

Answer:

$$P(X > 20) = \int_{20}^{\infty} f(x)dx = -e^{-x/20}\Big|_{20}^{\infty} = e^{-1} = 0.3679$$

Ex 3. X has a pdf $f(x) = 2x, \quad 0 < x < 1$. Find μ and σ^2.

Answer:

$$\mu = E(X) = \int_0^1 x f(x) dx = \int_0^1 2x^2 dx = \left(\frac{2}{3}x^3\right)\Big|_0^1 = \frac{2}{3}$$
$$E\left(X^2\right) = \int_0^1 x^2 f(x) dx = \int_0^1 2x^3 dx = \left(\frac{1}{2}x^4\right)\Big|_0^1 = \frac{1}{2}$$
$$\sigma^2 = E\left(X^2\right) - \{E(X)\}^2 = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}$$

Ex 4. X has a pdf $f(x) = \dfrac{3x^2}{4^3} e^{-(x/4)^3}, \quad 0 \le x < \infty$. Find $\pi_{0.9}$, i.e., the 90th percentile.

Answer:

$$\text{cdf } F(x) = \int_0^x f(t) dt = \int_0^x \frac{3t^2}{4^3} e^{-(t/4)^3} dt = -e^{-\frac{t^3}{4^3}}\Big|_0^x = 1 - e^{-(x/4)^3}, \quad 0 \le x < \infty$$

Since $F\left(\pi_{0.9}\right) = 0.9$, we have

$$1 - e^{-(x/4)^3} = 0.9$$
$$\ln(0.1) = -\left(\frac{x}{4}\right)^3$$
$$-\ln(0.1) = \left(\frac{x}{4}\right)^3$$

$$\therefore\ x = 4 \cdot \sqrt[3]{-\ln(0.1)} = 5.2820$$

Uniform distribution

Definition 2. $X \sim$ *uniform* $[a, b]$
X has a uniform distribution in the interval $[a, b]$, then
$\Leftrightarrow \quad f(x) = \dfrac{1}{b-a}, \quad a \le x \le b$
$\Leftrightarrow \quad M(t) = \dfrac{e^{tb} - e^{ta}}{t(b-a)}, \quad t \ne 0$

Properties 2.

$$\mu = \frac{a+b}{2}, \quad \sigma^2 = \frac{(b-a)^2}{12}$$

Proof.

$$M(t) = E\left(e^{tX}\right) = \int_a^b e^{tx} f(x) dx = \frac{1}{t(b-a)} e^{tx}\Big|_{x=a}^{x=b} = \frac{e^{tb} - e^{ta}}{t(b-a)}$$
$$E(X) = \int_a^b x f(x) dx = \frac{1}{2(b-a)} x^2\Big|_{x=a}^{x=b} = \frac{a+b}{2}$$

$$
E\left(X^2\right) = \int_a^b x^2 f(x)dx = \frac{1}{3(b-a)}\, x^3\Big|_{x=a}^{x=b} = \frac{b^2+ab+a^2}{3}
$$

$$
\therefore\ \sigma^2 = E\left(X^2\right) - \{E(X)\}^2 = \frac{b^2+ab+a^2}{3} - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}
$$

□

Exponential distribution

Definition 3. $X \sim \mathit{exponential}(\theta)$, $\theta > 0$

X has an exponential distribution with a parameter θ, $\theta > 0$, then

$\Leftrightarrow \quad f(x) = \dfrac{1}{\theta} e^{-x/\theta}, \quad 0 \le x < \infty$

$\Leftrightarrow \quad M(t) = (1-\theta t)^{-1}, \quad t < \dfrac{1}{\theta}$

Properties 3.

$$
\mu = \theta, \quad \sigma^2 = \theta^2
$$

Definition 4. $X \sim \Gamma(\alpha, \theta)$, $\alpha > 0$, $\theta > 0$

X has a gamma distribution with parameters α, θ, $\alpha > 0$, $\theta > 0$, then

$\Leftrightarrow \quad f(x) = \dfrac{1}{\Gamma(\alpha)\theta^\alpha} x^{\alpha-1} e^{-x/\theta}, \quad 0 \le x < \infty$

$\Leftrightarrow \quad M(t) = (1-\theta t)^{-\alpha}, \quad t < \dfrac{1}{\theta}$

Properties 4.

$$
\mu = \alpha\theta, \quad \sigma^2 = \alpha\theta^2
$$

Definition 5. $X \sim \chi^2(r)$, $r = 1, 2, \ldots$

X has a chi-square distribution with a parameter (=df) r, $r = 1, 2, \ldots$, then

$\Leftrightarrow \quad f(x) = \dfrac{1}{\Gamma(r/2)2^{r/2}} x^{r/2-1} e^{-x/2}, \quad 0 \le x < \infty$

$\Leftrightarrow \quad M(t) = (1-2t)^{-r/2}, \quad t < \dfrac{1}{2}$

Properties 5.

$$
\mu = r, \quad \sigma^2 = 2r
$$

One note:

- $Exp(\theta) = \Gamma(1, \theta)$ and $\chi^2(r) = \Gamma\left(\frac{r}{2}, 2\right)$
- $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx = (\alpha - 1)!$ for positive integer α

Proof. For $X \sim \Gamma(\alpha, \theta)$

$$
\begin{aligned}
M(t) = E\left(e^{tX}\right) &= \int_0^\infty e^{tx} \frac{1}{\Gamma(\alpha)\theta^\alpha} x^{\alpha-1} e^{-x/\theta} dx \\
&= \frac{1}{\Gamma(\alpha)\theta^\alpha} \frac{\Gamma(\alpha)}{\left(\frac{1}{\theta} - t\right)^\alpha} \int_0^\infty \frac{1}{\Gamma(\alpha)\left(\frac{1}{\theta} - t\right)^{-\alpha}} x^{\alpha-1} exp\left[-\left\{\frac{x}{\left(\frac{1}{\theta} - t\right)^{-1}}\right\}\right] dx
\end{aligned}
$$

The integrand part is the pdf of $\Gamma\left[\alpha, \left(\frac{1}{\theta} - t\right)^{-1}\right]$, thus the whole integral is one.

$$
M(t) = \frac{1}{\Gamma(\alpha)\theta^\alpha} \frac{\Gamma(\alpha)}{\left(\frac{1}{\theta} - t\right)^\alpha} = (1 - \theta t)^{-\alpha}, \quad for\,(1 - \theta t) > 0
$$

$\square$

Properties 6.

1. *$X_1, \ldots, X_n$; independent $X_i \sim \Gamma(\alpha_i, \theta)$, $i = 1, \ldots, n \Rightarrow \sum_{i=1}^n X_i \sim \Gamma\left(\sum_{i=1}^n \alpha_i, \theta\right)$*
2. *$X_1, \ldots, X_n$; iid $Exp(\theta) = \Gamma(1, \theta) \Rightarrow \sum_{i=1}^n X_i \sim \Gamma(n, \theta)$*
3. *$X_1, \ldots, X_n$; iid $\chi^2(r_i) = \Gamma\left(\frac{r_i}{2}, 2\right)$, $i = 1, \ldots, n \Rightarrow \sum_{i=1}^n X_i \sim \chi^2\left(\sum_{i=1}^n r_i\right)$*

Proof.

$$
\begin{aligned}
M_{\sum X_i}(t) = E\left(e^{t \sum X_i}\right) &= E\left(e^{tX_1} \cdots e^{tX_n}\right) \\
&= \prod_{i=1}^n E\left(e^{tX_i}\right) \qquad \text{(Why?)} \\
&= \prod_{i=1}^n M_{X_i}(t) \\
&= \prod_{i=1}^n (1 - \theta t)^{-\alpha_i} \qquad X_i \sim \Gamma(\alpha_i, \theta) \\
&= (1 - \theta t)^{-\sum_1^n \alpha_i} \qquad \sim \text{mgf of } \Gamma\left(\sum_1^n \alpha_i, \theta\right)
\end{aligned}
$$

By the uniqueness of the mgf, we have

$$
\sum_{i=1}^n X_i \sim \Gamma\left(\sum_1^n \alpha_i, \theta\right)
$$

$\square$

Properties 7.

1. $X \sim \Gamma(\alpha, \theta) \quad \Rightarrow \quad \dfrac{X}{\theta} \sim \Gamma(\alpha, 1)$. *This also means* $X = \theta Z, \quad Z \sim \Gamma(\alpha, 1)$
2. $X \sim \Gamma(\alpha, \theta) \quad \Rightarrow \quad cX \sim \Gamma(\alpha, c\theta)$
3. $X \sim \Gamma(\alpha, \theta) \quad \Rightarrow \quad \dfrac{2X}{\theta} \sim \Gamma(\alpha, 2) = \chi^2(2\alpha)$
4. $X \sim Exp(\theta) \quad \Rightarrow \quad \dfrac{2X}{\theta} \sim \Gamma(1, 2) = \chi^2(2)$

Ex 5. Suppose the lifetime of a certain electronic component has an exponential distribution with a mean of 500 hours. Let $X =$ the time to failure. Find (a) $P(X > 600)$, (b) $P(X > 900|X > 300)$.

Answer:

$$P(X > t) = \int_t^\infty f(x)dx = \int_t^\infty \frac{1}{500} e^{-x/500} dx = -e^{x/500}\Big|_{x=t}^{x=\infty} = e^{-t/500}$$
$$P(X > 600) = e^{-600/500} = e^{-6/5} = 0.3012$$
$$P(X > 900|X > 300) = \frac{P(X > 900)}{P(X > 300)} = \frac{e^{-900/500}}{e^{-300/500}} = e^{-6/5} = 0.3012$$

This is called the "memoryless" property of the exponential distribution. It says that the probability of a brand new one lasting over 600 hours is the same as that of the one that has already lasted over 300 hours. Of course, this is untrue in practice.

Definition 6. $X \sim Normal\left(\mu, \sigma^2\right)$
X has a normal distribution with parameters μ (=mean) and σ^2 (=variance), then

$$\Leftrightarrow \quad f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty$$
$$\Leftrightarrow \quad M(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)$$

Facts: $\displaystyle\int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2} dz = \sqrt{2\pi}$ and $\displaystyle 2\int_0^\infty e^{-\frac{1}{2}z^2} dz = \sqrt{2\pi}$

Proof. $M(t) = \exp\left(\mu t + \dfrac{\sigma^2 t^2}{2}\right)$
Let's first derive the mgf of $Z \sim N(0, 1)$.

$$\begin{aligned}
M_Z(t) = E\left(e^{tZ}\right) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{tz} e^{-\frac{1}{2}z^2} dz \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-t)^2} e^{\frac{1}{2}t^2} dz \\
&= e^{\frac{1}{2}t^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-t)^2} dz \\
&= e^{\frac{1}{2}t^2} \quad \because \text{The integrand part is pdf of } N(t, 1) \text{, i.e., the whole integral} = 1.
\end{aligned}$$

Next, note that $X \sim N\left(\mu, \sigma^2\right) \quad \Rightarrow \quad X = \sigma Z + \mu$, where $Z \sim N(0,1)$

$$M_X(t) = E\left(e^{tX}\right) = E\left(e^{t\sigma z + t\mu}\right) = e^{t\mu} \cdot M_Z(t\sigma) = e^{t\mu} e^{\frac{1}{2}\sigma^2 t^2} = e^{\mu t + \frac{1}{2}t^2}$$

□

Properties 8. $X \sim N\left(\mu, \sigma^2\right) \quad \Rightarrow \quad E(X) = \mu$, $Var(X) = \sigma^2$

Proof. $X \sim N\left(\mu, \sigma^2\right) \quad \Rightarrow \quad X = \sigma Z + \mu$, where $Z \sim N(0,1)$

So, $E(X) = \sigma E(Z) + \mu$ and $Var(X) = \sigma^2 Var(Z)$

$$\begin{aligned}
E(Z) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} z e^{-\frac{1}{2}z^2} dz = -\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \Big|_{-\infty}^{\infty} = 0 \\
Var(Z) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} z^2 e^{\frac{1}{2}t^2} dz - \{E(Z)\}^2 \\
&= \frac{2}{\sqrt{2\pi}} \int_0^{\infty} z d\left(-e^{-\frac{1}{2}z^2}\right) \\
&= \frac{2}{\sqrt{2\pi}} \left\{ -z e^{-\frac{1}{2}z^2} \Big|_0^{\infty} + \int_0^{\infty} e^{-\frac{1}{2}z^2} dz \right\} \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2} dz = 1
\end{aligned}$$

□

Properties 9.

1. $X \sim N\left(\mu, \sigma^2\right) \quad \Rightarrow \quad aX + b \sim N\left(a\mu + b,\, a^2\sigma^2\right)$
2. $X_1, \ldots, X_n$; *independent* $N\left(\mu_i, \sigma_i^2\right), i = 1, \ldots, n \quad \Rightarrow \quad \sum_{i=1}^{n} a_i X_i \sim \left(\sum a_i \mu_i + \sum a_i^2 \sigma_i^2\right)$

Proof.

$$\begin{aligned}
X \sim N\left(\mu, \sigma^2\right) &\Longleftrightarrow X \sim \sigma Z + \mu, \quad Z \sim N(0,1) \\
&\Longleftrightarrow aX + b \sim (a\sigma Z) + (a\mu + b), \quad Z \sim N(0,1) \\
&\Longleftrightarrow aX + b \sim N\left(a\mu + b, a^2\sigma^2\right)
\end{aligned}$$

$$
\begin{aligned}
\text{mgf } M_{\sum a_i X_i}(t) &= E\left\{\exp\left(\sum a_i X_i\right)\right\} = E\left(e^{ta_1X_1}\cdots e^{ta_nX_n}\right) \\
&= \prod_{i=1}^{n} E\left(e^{ta_iX_i}\right) \qquad (\because \text{ independent}) \\
&= \prod_{i=1}^{n} M_{X_i}\left(ta_i\right) \\
&= \prod_{i=1}^{n} \exp\left[ta_i\mu_i + \frac{\sigma_i^2}{2}\left(ta_i\right)^2\right], \quad X_i \sim N\left(\mu_i, \sigma_i^2\right) \\
&= \exp\left[t\left(\sum a_i\mu_i\right) + \frac{t^2}{2}\left(\sum a_i^2\sigma_i^2\right)\right] \quad \sim \text{ mgf of } N\left(\sum a_i\mu_i, \sum a_i^2\sigma_i^2\right)
\end{aligned}
$$

By the uniqueness of mgf, $\sum_{i=1}^{n} a_i X_i \sim N\left(\sum a_i\mu_i, \sum a_i^2\sigma_i^2\right)$ □

One note: Is $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ $(-\infty < x < \infty)$ really a pdf?

Answer: Let $Z = \frac{X-\mu}{\sigma}$, then $I = \int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$,

$$
I^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}}dx \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}dy = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{x^2+y^2}{2}}dx\,dy
$$

Let's use polar transformation, i.e., $x = r\cos\theta$ and $y = r\sin\theta$, and $\begin{vmatrix} \cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta \end{vmatrix} = r$

$$
I^2 = \frac{1}{2\pi}\int_0^{2\pi}\int_0^{\infty} e^{-\frac{r^2}{2}} r\,dr\,d\theta = \frac{1}{2\pi}\int_0^{2\pi} d\theta = 1
$$

$$
\therefore \quad I = 1 \quad (\because I > 0)
$$

□

Theorem 1. $Z \sim N(0,1) \quad \Rightarrow \quad Z^2 \sim \chi^2(1)$

Proof. (By cdf method) Let $Y = Z^2$

$$
\begin{aligned}
cdf\ \ G(y) = P(Y \le y) = P(-\sqrt{y} \le z \le \sqrt{y}) &= \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = 2\int_0^{\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz \\
&= \int_0^y \frac{1}{\sqrt{2\pi t}} e^{-\frac{t}{2}} dt \qquad (\text{Let } \ z = \sqrt{t}) \\
pdf\ \ g(y) = G'(y) = \frac{1}{\sqrt{2\pi y}} e^{-\frac{y}{2}},\ y > 0 \quad &\sim \Gamma\left(\frac{1}{2}, 2\right) = \chi^2(1), \quad \text{note also } \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}
\end{aligned}
$$

$\square$

Theorem 2. ***Fundamental results about random sample from normal population***

Let $X_1, X_2, \ldots, X_n$: i.i.d. from $N(\mu, \sigma^2)$. Then,

1. $\bar{X} \sim N\left(\mu, \dfrac{\sigma^2}{n}\right)$, *i.e.,* $\dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$
2. $\bar{X}$ *&* $S^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=1}^n \left(X_i - \bar{X}\right)^2$ *are independent.*
3. $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$

Proof. First, note that it is sufficient to show that $\bar{X}$ & $\left(X_1 - \bar{X}, \ldots, X_n - \bar{X}\right)$ are indepdendent because $S^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=1}^n \left(X_i - \bar{X}\right)^2$ is a function of $\left(X_1 - \bar{X}, \ldots, X_n - \bar{X}\right)$.

The mgf of $\bar{X}$ & $\left(X_1 - \bar{X}, \ldots, X_n - \bar{X}\right)$ is given by

$$
\begin{aligned}
M(s, t_1, \cdots, t_n) &= E\left[\exp\left\{s\bar{X} + t_1\left(X_1 - \bar{X}\right) + \cdots + t_n\left(X_n - \bar{X}\right)\right\}\right] \\
&= E\left[\exp\left\{\sum_1^n \left(t_i + \frac{s - t_1 - \cdots - t_n}{n}\right) X_i\right\}\right] \\
&= \prod_{i=1}^n E\left[\exp\left\{\left(t_i + \frac{s - t_1 - \cdots - t_n}{n}\right) X_i\right\}\right] \qquad \because \text{indepndent} \\
&= \prod_{i=1}^n M_{X_i}\left(t_i + \frac{s - \sum t_i}{n}\right) \qquad \because \text{definition} \\
&= \prod_{i=1}^n \exp\left\{\mu\left(t_i + \frac{s - \sum t_i}{n}\right) + \frac{\sigma^2}{2}\left(t_i + \frac{s - \sum t_i}{n}\right)^2\right\} \qquad \because X_i \sim N\left(\mu, \sigma^2\right)
\end{aligned}
$$

Here, let's first sort out the last part:

$$\begin{aligned}\left(t_i + \frac{s - \sum t_i}{n}\right)^2 &= t_i^2 + \frac{2t_i(s - \sum t_i)}{n} + \frac{(s - \sum t_i)^2}{n^2} \\ &= \frac{s^2}{n^2} - \frac{2\sum t_i s}{n^2} + \frac{2t_i s}{n} + t_i^2 - \frac{2t_i \sum t_i}{n} + \frac{(\sum t_i)^2}{n^2}\end{aligned}$$

Putting all these together, look at the parts inside "exp" part (after multiplication):

$$\begin{aligned}\mu \sum t_i + \mu s - \mu \sum t_i + \frac{\sigma^2}{2}\sum t_i^2 + \frac{\sigma^2(s - \sum t_i)\sum t_i}{n} + \frac{\sigma^2\{s^2 - 2s\sum t_i + (\sum t_i)^2\}}{2n} \\ = \mu s + \frac{s^2}{2}\frac{\sigma^2}{n} + \frac{\sigma^2}{2}\left\{\sum t_i^2 - \frac{(\sum t_i)^2}{n}\right\}\end{aligned}$$

So we have

$$M(s, t_1, \cdots, t_n) = \exp\left(\mu s + \frac{s^2}{2}\frac{\sigma^2}{n}\right) \cdot \exp\left[\frac{\sigma^2}{2}\left\{\sum t_i^2 - \frac{(\sum t_i)^2}{n}\right\}\right]$$

This means

$$M_{\bar{X}} = M(s, 0, \cdots, 0) = \exp\left(\mu s + \frac{s^2}{2}\frac{\sigma^2}{n}\right)$$

$$M_{\left(X_1 - \bar{X}, \cdots, X_n - \bar{X}\right)} = M(0, t_1, \cdots, t_n) = \exp\left[\frac{\sigma^2}{2}\left\{\sum t_i^2 - \frac{(\sum t_i)^2}{n}\right\}\right]$$

That is,

$$M_{\bar{X}, \left(X_1 - \bar{X}, \cdots, X_n - \bar{X}\right)}(s, t_1, \cdots, t_n) = M(s, 0, \cdots, 0) \cdot M(0, t_1, \cdots, t_n)$$

Therefore, $\bar{X}$ and $\left(X_1 - \bar{X}, \cdots, X_n - \bar{X}\right)$ are independent.

Thus, $\bar{X}$ and $S^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$ are independent. $\square$

MLE: Maximum Likelihood Estimate in the Continuous Case

Ex 6. $X_1, X_2, \ldots, X_n$: iid from $Exp(\theta)$. Find the mle of θ.

Answer:

$$
\begin{aligned}
L(\theta) &= \prod_{i=1}^{n} f(x_i;\theta) = \prod_{i=1}^{n}\left(\frac{1}{\theta}e^{-x_i/\theta}\right), \quad \theta \in \Omega \\
&= \left(\frac{1}{\theta}\right)^n e^{-\sum x_i/\theta} \\
\ln L(\theta) &= -n\ln\theta - \sum_{i=1}^{n}\frac{x_i}{\theta} \quad \text{(log-likelihood function)}
\end{aligned}
$$

To find θ that maximizes the log likelihood, differentiate this wrt θ and set it equal to zero, we get

$$
\begin{aligned}
&\frac{\partial}{\partial\theta}\ln L(\theta) = -\frac{n}{\theta} + \frac{\sum x_i}{\theta^2} = 0 \\
&\hat{\theta} = \frac{\sum x_i}{n} = \bar{x}
\end{aligned}
$$

To make sure that this indeed makes the log likelihood maximum, do the second derivative test. Here, we have

$$
\frac{\partial^2}{\partial^2\theta}\ln L(\theta) = \frac{n}{\theta^2} - 2\frac{(\sum x_i)}{\theta^3}\Big|_{\hat{\theta}} = -\frac{n}{\bar{x}^2} < 0
$$

This means that our solution $\hat{\theta} = \bar{X}$ is indeed the mle.

Ex 7. $X_1, X_2, \ldots, X_n$: iid from $N(\mu = \theta, \sigma^2)$. Find the mle of θ.

Answer:

$$
\begin{aligned}
L(\theta) &= \prod_{i=1}^{n} f(x_i;\theta) = \prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x_i-\theta)^2}{2\sigma^2}}, \quad -\infty < \theta < \infty \\
&= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\sum (x_i-\theta)^2/2\sigma^2} \\
\ln L(\theta) &= -n\ln\left(\sqrt{2\pi}\sigma\right) - \sum_{i=1}^{n}\frac{(x_i-\theta)^2}{2\sigma^2}
\end{aligned}
$$

Differentiate this wrt θ and set it equal to zero, we get

$$\frac{\partial}{\partial\theta}\ln L(\theta)=\sum_{i=1}^{n}\frac{(x_i-\theta)}{\sigma^2}=0$$
$$\hat{\theta}=\frac{\sum x_i}{n}=\bar{x}$$

Check the second derivative, we get

$$\frac{\partial^2}{\partial^2\theta}\ln L(\theta)=-\frac{n}{\sigma^2}<0$$

This means that our solution $\hat{\theta}=\bar{X}$ is indeed the mle.

Ex 8. $X_1, X_2, \ldots, X_n$: iid from $N(\theta_1, \theta_2)$. Find the mle of θ_1 and θ_2.

Answer:

$$\begin{aligned}L(\theta_1,\theta_2)&=\prod_{i=1}^{n} f(x_i;\theta_1,\theta_2)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\theta_2}}e^{-\frac{(x_i-\theta_1)^2}{2\theta_2}}, \quad -\infty<\theta_1<\infty;\ 0<\theta_2\\ &=\left(\frac{1}{\sqrt{2\pi\theta_2}}\right)^n e^{-\sum(x_i-\theta_1)^2/2\theta_2}\\ \ln L(\theta_1,\theta_2)&=-\frac{n}{2}\ln(2\pi\theta_2)-\sum_{i=1}^{n}\frac{(x_i-\theta_1)^2}{2\theta_2}\end{aligned}$$

Because of the two parameters, we differentiate this wrt θ_1, θ_2 and set them equal to zero, we get

$$\frac{\partial}{\partial\theta_1}\ln L(\theta_1,\theta_2)=\frac{1}{\theta_2}\sum(x_i-\theta_1)=0 \quad\Rightarrow\quad \hat{\theta_1}=\bar{x}$$
$$\frac{\partial}{\partial\theta_2}\ln L(\theta_1,\theta_2)=-\frac{n}{2\theta_2}+\frac{1}{2\theta_2^2}\sum(x_i-\theta_1)^2=0 \quad\Rightarrow\quad \hat{\theta_2}=\frac{1}{n}\sum(x_i-\bar{x})^2$$

Now, to check the second derivatives separately and show that they are negative at $\hat{\theta_1}$ and $\hat{\theta_2}$ is NOT right, but here they are anyway:

$$\frac{\partial^2}{\partial\theta_1^2}\ln L(\theta_1,\theta_2)=-\frac{n}{\theta_2}\Big|_{\hat{\theta_1},\hat{\theta_2}}<0$$
$$\frac{\partial^2}{\partial\theta_2^2}\ln L(\theta_1,\theta_2)=\frac{n}{2\theta_2^2}-\frac{1}{\theta_2^3}\sum(x_i-\theta_1)^2\Big|_{\hat{\theta_1},\hat{\theta_2}}=\frac{1}{2\theta_2^3}\left\{n\hat{\theta_2}-2\sum(x_i-\bar{x})^2\right\}<0$$

BUT this is NOT right. The correct way of showing that the log likelihood indeed maximizes at $\hat{\theta_1}$ and $\hat{\theta_2}$ is by showing the following 2×2 matrix is negative definite. Here is the right way to show this.

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

$$\begin{aligned}
A &= \frac{\partial^2}{\partial \theta_1^2} \ln L(\theta_1, \theta_2) = -\frac{n}{\theta_2} \\
B = C &= \frac{\partial^2}{\partial \theta_1 \theta_2} \ln L(\theta_1, \theta_2) = -\frac{1}{\theta_2^2} \sum (x_i - \theta_1) \\
D &= \frac{\partial^2}{\partial \theta_2^2} \ln L(\theta_1, \theta_2) = \frac{n}{2\theta_2^2} - \frac{1}{\theta_2^3} \sum (x_i - \theta_1)^2
\end{aligned}$$

We're done if we show the determinant of this matrix is negative. Here is the last part.

$$\begin{aligned}
& -\frac{n}{\theta_2} \left\{ \frac{n}{2\theta_2^2} - \frac{1}{\theta_2^3} \sum (x_i - \theta_1)^2 \right\} - \frac{1}{\theta_2^4} \left\{ \sum (x_i - \theta_1) \right\}^2 \\
&= \frac{n}{2\theta_2^4} \left[-n\theta_2 + 2 \sum (x_i - \theta_1)^2 - 2n \left\{ \sum (x_i - \theta_1) \right\}^2 \right] \Big|_{\hat{\theta}_1, \hat{\theta}_2} \\
&= \frac{n}{2\hat{\theta_2}^4} \left[-\sum (x_i - \bar{x})^2 + 2 \sum (x_i - \bar{x})^2 - 2n \left\{ \sum \left(x_i - \hat{\theta}_1 \right) \right\}^2 \right] \\
&= \frac{n}{2\hat{\theta_2}^4} \left[\sum (x_i - \bar{x})^2 - 2n \left\{ \sum (x_i - \bar{x})^2 + \sum \sum\nolimits_{i<j} (x_i - \bar{x})(x_j - \bar{x}) \right\} \right] \quad < 0
\end{aligned}$$

$\square$

CLT: The Central Limit Theorem

Theorem 3. *Let $X_1, \ldots, X_n$ be random samples from a distribution with mean μ and variance σ^2. Then,*

$$W = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} = \frac{\sqrt{n} \left(\bar{X} - \mu \right)}{\sigma} = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \quad \xrightarrow{d} \quad N(0,1)$$

Proof.

$$
\begin{aligned}
M_W(t;n) = E\left(e^{tW}\right) = E\left\{ e^{t\dfrac{\sqrt{n}(\bar{X}-\mu)}{\sigma}} \right\} &= E\left\{\exp\left(\frac{t\sqrt{n}}{\sigma}\bar{X}\right)\right\}\cdot \exp\left(-\frac{\mu}{\sigma}\sqrt{n}t\right) \\
&= \exp\left(-\frac{\mu}{\sigma}\sqrt{n}t\right)\prod_{i=1}^{n} E\left\{\exp\left(\frac{t}{\sqrt{n}\sigma}X_i\right)\right\} \quad (\because \text{ independent}) \\
&= \exp\left(-\frac{\mu}{\sigma}\sqrt{n}t\right)\left\{M_{X_i}\left(\frac{t}{\sqrt{n}\sigma}\right)\right\}^n \quad (\because \text{ identical}) \\
&= \left\{\exp\left(-\frac{\mu}{\sqrt{n}\sigma}t\right) M_{X_i}\left(\frac{t}{\sqrt{n}\sigma}\right)\right\}^n \\
&= \left\{m\left(\frac{t}{\sqrt{n}\sigma}\right)\right\}^n, \quad \text{where } m(t) = e^{-\mu t}M(t) \\
&= \left[m(0) + m'(0)\cdot\left(\frac{t}{\sqrt{n}\sigma}\right) + \frac{m''(0)}{2}\cdot\left(\frac{t}{\sqrt{n}\sigma}\right)^2 + \frac{o(1)}{n} + \cdots\right]^n \\
&= \left[1 + E(X_1-\mu)\cdot\left(\frac{t}{\sqrt{n}\sigma}\right) + \frac{E(X_1-\mu)^2}{2}\cdot\left(\frac{t}{\sqrt{n}\sigma}\right)^2 + \frac{o(1)}{n} + \cdots\right]^n \qquad (1)
\end{aligned}
$$

To see the last move, look at $m(t) = e^{-\mu t}M(t) = E\left\{e^{t(X_1-\mu)}\right\} = \text{mgf of } (X_1-\mu)$

$$
(1) = \left[1 + \frac{1}{2}\frac{t^2}{n} + \frac{o(1)}{n}\right]^n \quad \xrightarrow{d} \quad \exp\left(\frac{1}{2}t^2\right) = \text{mgf of } N(0,1)
$$

□

Ex 9. Let $\bar{X}$ be the mean of a random sample of $n = 25$ from a population with a mean of 15 and a variance of 4. Find $P\left(14.4 < \bar{X} < 15.6\right)$.

Answer: $\bar{X} \dot{\sim} N\left(mean = \mu = 15;\, var = \sigma^2/n = 4/25 = 0.16\right)$

$$
P\left(14.4 < \bar{X} < 15.6\right) = P\left(\frac{14.4-15}{0.4} < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < \frac{15.6-15}{0.4}\right) \approx \Phi(1.5) - \Phi(-1.5) = 0.8664
$$

Ex 10. $X_1, \ldots, X_{20}$: random samples from uniform $(0,1)$. Let $Y = X_1 + \cdots + X_{20}$. Find (a) $P(Y \le 9.1)$, (b) $P(8.5 \le Y \le 11.7)$.

Answer: $E(X_i) = \frac{1}{2}$, $Var(X_i) = \frac{1}{12}$. Also, $E(Y) = 10$, $Var(Y) = \frac{20}{12}$ (Why?). This means

$Y \dot{\sim} N\,(mean = 10;\, var = 20/12)$

$$P\,(Y \le 9.1) = P\left(\frac{Y-\mu_Y}{\sigma_Y} < \frac{9.1-10}{\sqrt{20/12}}\right) \approx \Phi(-0.697) = 0.2423$$
$$P\,(8.5 \le Y \le 11.7) = P\left(\frac{8.5-10}{\sqrt{20/12}} \le \frac{Y-\mu_Y}{\sigma_Y} \le \frac{11.7-10}{\sqrt{20/12}}\right) \approx \Phi(1.317) - \Phi(-1.162) = 0.7835$$

Facts: *Normal approximation to various distributions*

- $X \sim binomial\,(n,p)$, $0<p<1$, then $\dfrac{X-np}{\sqrt{npq}} \quad \xrightarrow{d} \quad N(0,1)$
- $X \sim Poisson\,(\lambda)$, $\lambda > 0$, then $\dfrac{X-\lambda}{\sqrt{\lambda}} \quad \xrightarrow{d} \quad N(0,1)$
- $X \sim \chi^2\,(r)$, $r>0$, then $\dfrac{X-r}{\sqrt{2r}} \quad \xrightarrow{d} \quad N(0,1)$

Ex 11. $X \sim binomial\left(36, \frac{1}{2}\right)$. Find $P(12 < X \le 18)$.

Answer:

$$\begin{aligned} P\,(12 < X \le 18) &= P\,(12.5 \le X \le 18.5) \quad (\because \text{continuity correction}) \\ &= P\left(\frac{12.5-18}{3} \le \frac{X-\mu_X}{\sigma_X} \le \frac{18.5-18}{3}\right) \\ &\approx \Phi(0.167) - \Phi(-1.833) = 0.5329 \end{aligned}$$