Lab 3 : Theory

This part is overridden by the following parts !

1) Parameters of Gamma distr. are

- α : shape parameter
- β : rate parameter

NOTE: $\alpha, \beta > 0$

PDF of Gamma(α, β) is

$$\text{Gamma}(\alpha, \beta)(x) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$$

χ^2 distribution is a special case of Gamma, wherein $\alpha = k/2$ & $\beta = 1/2$

PDF of $\chi^2(k)$ is

$$\chi^2(k)(x) = \frac{1}{2^{k/2}\Gamma(k/2)} x^{k/2 - 1} e^{-x/2}$$

NOTE: Gamma can be parametrised in 2 ways :

NOTE: All params here are positive

(1) shape parameter α, rate parameter β

(2) shape parameter α, scale parameter θ

Here, note that $\beta = 1/\theta$

In the question, Gamma is parametrised in the 2nd way. Hence, PDF of Gamma is

~~Gamma(α, θ) =~~

$$\text{Gamma}(\alpha, \theta)(x) = \frac{1}{\theta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-x/\theta}$$

if $x > 0$, (else Gamma(α, θ)(x) = 0)

χ^2 distribution is a special case of Gamma wherein ~~$\alpha = \theta$~~

$\alpha = k/2$, $\theta = \cancel{1/2}\ 2$

Hence, PDF of $\chi^2(k)$ is

$$\chi^2(k)(x) = \frac{1}{2^{k/2}\Gamma(k/2)} x^{k/2-1} e^{-x/2}$$

Now, χ^2 has additivity i.e.

$$\chi^2(k_1) + \chi^2(k_2) \dots + \chi^2(k_n) = \chi^2(k_1 + \dots + k_n)$$

(provided each distribution is independent).

$$\therefore \sum_{i=1}^{N} \chi^2(k) = \chi^2(Nk)$$

Hence, consider $\frac{1}{N}\chi^2(Nk)$

In ~~essence~~ essence $\frac{1}{N}\chi^2(Nk)$ is the push-forward measure

$$g_* \chi^2(Nk) \quad \text{where } g: \mathbb{R} \to \mathbb{R},\ x \mapsto \frac{x}{N}.$$

Hence, for any $x > 0$,

$$\underbrace{g_* \chi^2(Nk)}_{\downarrow}((-\infty, x])$$

This is the CDF, which gives the probability mass of a subset and is hence a measure (probability measure to be exact)

$$= \chi^2(Nk)\left(\left\{x_0 \mid \frac{x_0}{N} \in (-\infty, x],\ x_0 \in \mathbb{R}\right\}\right)$$

(by definition of pushforward)

$$= \chi^2(Nk)((-\infty, Nx])$$

Now, by the fundamental theorem of calculus, we have that if $F: \mathbb{R} \to [0,1]$ is the CDF of a probability distribution, then the corresponding PDF ~~is given by f~~ is given by

$$f(x) = \frac{dF(x)}{dx}$$

$$\therefore \frac{d\, g_* \chi^2(Nk)((-\infty, x])}{dx}$$

$$= \frac{d\, \chi^2(Nk)((-\infty, Nx])}{dNx}$$

$$= \chi^2(Nk)(Nx)$$

$$= \frac{1}{2^{Nk/2}\,\Gamma(Nk/2)} x^{Nk/2 - 1} e^{-Nx/2}$$

$$= \mathrm{Gamma}\left(\frac{Nk}{2}, \frac{2}{N}\right)(x)$$

NOTE: Alternatively, you could just use the CDF of χ^2 & Gamma, but in case you knew only PDF, you could do this

2) a) Let $g : \mathbb{R} \to \mathbb{R},\ x \mapsto x - \mu$

Then, for any interval $[a,b]$,

$g_* \mathrm{Norm}(\mu,\sigma)([a,b])$

$= \mathrm{Norm}(\mu,\sigma)(\{x \mid g(x) \in [a,b], x \in \mathbb{R}\})$

~~= Norm(μ, σ)~~ by definition of pushforward

$= \mathrm{Norm}(\mu,\sigma)(\{x \mid x - \mu \in [a,b], x \in \mathbb{R}\})$

$= \mathrm{Norm}(\mu,\sigma)([a+\mu,\ b+\mu])$

lower bound of x ... upper bound of x

... 1

Now note that $g_* \mathrm{Norm}(\mu,\sigma)$

$= \mathrm{Norm}(\mu,\sigma) - \mu$

(i.e. shifting the distribution to the left (toward the origin) by μ

$= \mathrm{Norm}(0,\sigma)$ (since shifting the disto. ~~mean~~ by k changes disto. mean by k also)

$\therefore$ 1 $\Rightarrow \mathrm{Norm}(0,\sigma)([a,b])$

$= \mathrm{Norm}(\mu,\sigma)([a+\mu,\ b+\mu])$

$\Rightarrow \mathrm{Norm}(0,\sigma)([a-\mu,\ b-\mu])$

$= \mathrm{Norm}(\mu,\sigma)([a,\ b])$

---

b) Let $h : \mathbb{R} \to \mathbb{R},\ x \mapsto \frac{x}{\sigma}$

Then, for any interval $[a,b]$,

$h_* \mathrm{Norm}(0,\sigma)([a,b])$

$= \mathrm{Norm}(0,\sigma)(\{x \mid h(x) \in [a,b], x \in \mathbb{R}\})$

~~by definition~~ of pushforward

$= \mathrm{Norm}(0,\sigma)(\{x \mid \frac{x}{\sigma} \in [a,b], x \in \mathbb{R}\})$

$= \mathrm{Norm}(0,\sigma)([\sigma a,\ \sigma b])$

... 2

Now, note that $h_* \text{Norm}(0,\sigma)$

$= \frac{1}{\sigma}\text{Norm}(0,\sigma) = \text{Norm}(0,1)$

$\therefore$ (2) $\Rightarrow \text{Norm}(0,1)([a,b])$

$= \text{Norm}(0,\sigma)([\sigma a, \sigma b])$

$\Rightarrow \text{Norm}(0,1)([\frac{a}{\sigma}, \frac{b}{\sigma}])$

$= \text{Norm}(0,\sigma)([a,b])$

---

c) Given h & g defined in (a) & (b), we have that

$\text{Norm}(0,1) = \cancel{g}\, h_*(g_* \text{Norm}(\mu,\sigma))$

~~= h*~~ Hence, using results of (a) & (b), we get...

$\text{Norm}(0,1)([\frac{a-\mu}{\sigma}, \frac{b-\mu}{\sigma}])$

$= \text{Norm}(0,1)([h(g(a)), h(g(b))])$

$= h_*(g_* \text{Norm}(\mu,\sigma))([h(g(a)), h(g(b))])$

* NOTE that ~~Norm~~ $g_* \text{Norm}(\mu,\sigma)$
$= \text{Norm}(0,\sigma)$ (as seen in (a))
Hence, from (b), we get:

$= \cancel{h_*}\, g_* \text{Norm}(\mu,\sigma)([g(a), g(b)])$

$= \text{Norm}(\mu,\sigma)([a,b])$

(from (a))

TIP: It might have been easier to relate to results of (a) & (b) if h & g were written out as expressions rather than symbolically
i.e. if $h(g(a))$ was written as $\frac{a-\mu}{\sigma}$
& ~~b~~ $g(a)$ as $a-\mu$
(same for b)

$\therefore \text{Norm}(0,1)([\frac{a-\mu}{\sigma}, \frac{b-\mu}{\sigma}])$

$= \text{Norm}(\mu,\sigma)([a,b])$

If Φ is the CDF of Norm(0,1), then

$\text{Norm}(0,1)([\frac{a-\mu}{\sigma}, \frac{b-\mu}{\sigma}]) =$

$$\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

---

d) The weak law of large numbers states that the spread of ~~n-sized~~ sample means of n-sized sample sets taken from a distribution $\mathbb{P}$ converges to an arbitrarily small interval around the mean of $\mathbb{P}$, as $n \to \infty$.

i.e. $$\lim_{n \to \infty} \overline{\mathbb{P}}_n([\mu(\mathbb{P}) - \varepsilon, \mu(\mathbb{P}) + \varepsilon]) = 1$$

$$\forall \varepsilon > 0$$

If $\mathbb{P}$ is the normal distribution ~~Norm~~ $\text{Norm}(\mu, \sigma)$, then

$$\overline{\mathbb{P}}_n = \text{Norm}\left(\mu, \frac{\sigma}{\sqrt{n}}\right), \quad \mu(\mathbb{P}) = \mu$$

This is given in the question, & obtained in lab 2's theory questions

Hence, consider the probability mass

$$\overline{\mathbb{P}}_n([\mu(\mathbb{P}) - \varepsilon, \mu(\mathbb{P}) + \varepsilon])$$

$$= \text{Norm}\left(\mu, \frac{\sigma}{\sqrt{n}}\right)([\mu - \varepsilon, \mu + \varepsilon]), \quad \varepsilon > 0$$

Let $\delta > 0$ s.t., for a fixed $\varepsilon > 0$,

$$\text{Norm}\left(\mu, \frac{\sigma}{\sqrt{n}}\right)([\mu - \varepsilon, \mu + \varepsilon]) > 1 - \delta$$

Now, note that from (c), we get

$$\text{Norm}\left(\mu, \frac{\sigma}{\sqrt{n}}\right)([\mu - \varepsilon, \mu + \varepsilon])$$

$$= \text{Norm}(0,1)\left(\left[\frac{(\mu - \varepsilon) - \mu}{\sigma / \sqrt{n}}, \frac{(\mu + \varepsilon) - \mu}{\sigma / \sqrt{n}}\right]\right)$$

$$= \text{Norm}(0,1)\left(\left[-\frac{\sqrt{n}\,\varepsilon}{\sigma}, \frac{\sqrt{n}\,\varepsilon}{\sigma}\right]\right)$$

$$= \Phi(+\sqrt{n}\,\varepsilon/\sigma) - \Phi(-\sqrt{n}\,\varepsilon/\sigma)$$

$$\therefore \Phi(\sqrt{n}\,\varepsilon/\sigma) - \Phi(-\sqrt{n}\,\varepsilon/\sigma) > 1 - \delta$$

Now, note that since ~~normal~~ $\text{Norm}(0,1)$ is symmetric about 0,

$$\Phi(\sqrt{n}\,\varepsilon/\sigma) - \Phi(-\sqrt{n}\,\varepsilon/\sigma)$$

$$= 1 - 2\Phi(-\sqrt{n}\,\varepsilon/\sigma)$$

$\therefore\ 1 - 2\,\Phi(-\sqrt{n}\,\varepsilon/\sigma) > 1 - \delta$

$\Rightarrow\ \Phi(-\sqrt{n}\,\varepsilon/\sigma) < \delta/2$

$\Rightarrow\ -\sqrt{n}\,\varepsilon/\sigma < \Phi^{-1}(\delta/2)$

(~~we can do~~ the above holds because any CDF is monotonic increasing, and CDF of a normal distribution is strictly monotonic increasing)

$\Rightarrow\ -\sqrt{n} < \sigma\,\Phi^{-1}(\delta/2)/\varepsilon$

$\Rightarrow\ \sqrt{n} > -\sigma\,\Phi^{-1}(\delta/2)/\varepsilon$

$\Rightarrow\ n > (\sigma\,\Phi^{-1}(\delta/2)/\varepsilon)^2 \quad \ldots (1)$

~~Now, since~~ ~~NOT NECESSARY~~

- ~~$\sigma > 0$~~
- ~~$\varepsilon > 0$~~
- ~~$\Phi^{-1}(\delta/2) \neq 0$ if $\delta \neq 1$~~

~~we have that if $\delta \neq 1$,~~

~~$(\sigma\,\Phi^{-1}(\delta/2)/\varepsilon)^2 > 0$~~

$\therefore$ ~~$\exists\, n > 0$ s.t.~~ For any n above a certain non-negative quantity, we have that for an arbitrary $\delta > 0$ and $\varepsilon > 0$,

$$\mathrm{Norm}(\mu, \tfrac{\sigma}{\sqrt{n}})([\mu - \varepsilon, \mu + \varepsilon]) > 1 - \delta$$

~~Furthermore, inequality (1) implies that for smaller values of δ, since Φ~~

Furthermore, since $\Phi^{-1}(\delta/2) \to -\infty$ as $\delta \to 0$, the inequality in (1) implies that ~~$n \to \infty$ as~~ as $\delta \to 0$, $n \to \infty$,

$$\Rightarrow \lim_{n\to\infty} \mathrm{Norm}(\mu, \tfrac{\sigma}{\sqrt{n}})([\mu - \varepsilon, \mu + \varepsilon]) = 1$$

Thus proving the weak law of large numbers for normal distribution.