14 רווח$TZ_1 5-8$

סיכום של ההתפלגות נורמית וההתפלגות הנורמלית הסטנדרדית

14.1 הגדרה. (מ"מ רציף נורמאלי) משתנה מקרי X מתפלג נורמאלי מוגדר להיות כך שצפיפותו נתון ע"י הנוסחאה

$$n(x,\mu,\sigma)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right),$$

כאשר μ הוא התוחלת ו- σ הסטיית התקן. מסמנים מ"מ נורמאלי ב $X\sim N(\mu,\sigma)$.

14.2 הגדרה. (מ"מ רציף נורמאלי סטנדרדי) משתנה מקרי Z מתפלג נורמאלי סטנדרדי מוגדר להיות מ"מ נורמאלי בעל תוחלת $\mu=0$ ו- סטיית התקן $\sigma=1$, כלומר המ"מ נורמאלי עם צפיפות נתון ע"י

$$n(z,\mu=0,\sigma=0)=\frac{1}{\sqrt{2\pi}}e^{-z^2/2} .$$

14.3 הגדרה. (התפלגות המצטברת של מ"מ נורמאלי סטנדרדי) התפלגות המצטברת של מ"מ נורמאלי סטנדרדי Z מוגדרת להיות

$$\Phi(z)=\int_{-\infty}^{z}dt\ \frac{1}{\sqrt{2\pi}}e^{-t^2/2}=\frac{1}{2}\left[1+\text{erf}\left(\frac{z}{\sqrt{2}}\right)\right] ,$$

כאשר

$$\text{erf}\,(z):=\frac{2}{\sqrt{\pi}}\int_0^z dt\ e^{-t^2} ,$$

(כלומר $\text{erf}\left(z/\sqrt{2}\right)=\frac{2}{\sqrt{\pi}}\int_0^{z/\sqrt{2}}dt\ e^{-t^2}=\frac{1}{\sqrt{2\pi}}\int_0^z dt\ e^{-t^2}$) והתכונה ש

$$\text{erf}(-z)=-\text{erf}(z)$$

נובע למסקנה

$$\Phi(-z)=1-\Phi(z) .$$

14.4 הגדרה. (הסתברות של מ"מ נורמאלי) ההסתברות שמ"מ נורמאלי X בעל תוחלת μ וסטיית התקן σ מקבל ערך פחות או שווה ל x_1 נתון ע"י

$$P(X\le x_1)=\Phi(z_1)$$

כאשר

$$z_1=\frac{x_1-\mu}{\sigma} .$$

ההסתברות שמ"מ נורמאלי X בעל תוחלת μ וסטיית התקן σ מקבל ערים בין x_1 ו- x_2 נתון ע"י

$$P(x_1\le X\le x_2)=\Phi(z_2)-\Phi(z_1)$$

כאשר

$$z_1=\frac{x_1-\mu}{\sigma} ,\qquad z_2=\frac{x_2-\mu}{\sigma} .$$

14.5 חוק. (קירוב נורמלי להתפלגות בינומית) יהי $X\sim\text{Bin}\,(n,p)$. אז עבור $n\ge 30$,

$$X\sim N\,(np,nq) .$$

תיקון רציפות:

$$P(a\le X\le b)$$

כמשתנה בדיד שווה ל

$$P(a-0.5<X<b+0.5)$$

כמשתנה רציף.

משפט הגבול המרכזי

14.6 חוק. (משפט הגבול המרכזי) יהי X משתנה מקרי בעל תוחלת μ וסטיית התקן σ, ויהי

$$x_1, \ldots, x_n$$

מדגם מקרי מתוך X. אזי,כאשר $n \geq 30$ או לכל n במקרה ש X מתפלג נורמלית, מתקיים

$$\sum_{i=1}^{n} X_i \sim N\left(n\mu, n\sigma^2\right),$$

או במילים אחרות, כאשר $\bar{X} := \frac{1}{n}\sum_{i=1}^{n} X_i$ הוא התוחלת של המדגם,

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) .$$

14.7 מסקנה. (משפט הגבול המרכזי) יהי $\bar{X}$ התוחלת של מדגם מקרי של אורך n מתוך אוכלוסיה בעל תוחלת μ ושונות σ^2. אזי,כאשר $n \geq 30$, ההתפלגות של המשתנה

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

היא ההתפלגות הנורמלית הסטנדרדית $n(z, 0, 1)$, קרי

$$Z \sim N(0, 1) .$$

רווחי סמך לתוחלת של מדגם מקרי: שונות ידועה

נניח שיש אוכלוסיה בעל תוחלת μ אינה ידועה ושונות σ^2 ידועה. עבור מדגם X כלשהו מתוך האוכלוסיה זו, לפי המשפט הגבול המרכזי, ההתפלגות של $\bar{X}$ מתפלג בקירוב נורמלי עם תוחלת $\mu_{\bar{X}} = \mu$ וסטיית התקן $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$.
השאלה היא:

נתון $\bar{X}$, מהו הטווח

$$a \leq \mu \leq b$$

כך כי יש הסתברות $(1 - \alpha)$ ל μ כן להיות נמצא בטווח זו, קרי

$$P(a \leq \mu \leq b) = 1 - \alpha$$

והסתברות α ל μ לא להיות נמצא בטווח זו, כלומר

$$P(\mu \notin [a, b]) = \alpha .$$

לדוגמה, מהו הטווח של μ כך שיש הסתברות של 0.95 (או 95%) להיות נמצא בו? הנה $1 - \alpha = 0.95$ אזי $\alpha = 0.05$. אנו מתבקשים למצוא a ו- b כך ש $P(a \leq \mu \leq b) = 0.95$. נגלה להלן [עיין משוואה (2*)] שהערכים הנדרשים הם

$$a = \bar{X} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} , \qquad b = \bar{X} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} ,$$

כאשר $\bar{X}$ הוא התוחלת המחושב של המדגם מקרי X, σ הסטיית התקן הידוע של האוכלוסיה כולה שממנה המדגם נלקח, n האורך של המדגם ו- $z_{1-\alpha/2}$ הוא הערך של המשתנה נורמלי סטנדרדי z המתאים ל $\bar{X}$ [עיין נוסחאה (1*)

להלן], מוגדר כך שהשטח התחום בטווח $[-z_{1-\alpha/2}, z_{1-\alpha/2}]$ של הגרף [עיין אייור להלן] שווה ל $(1-\alpha)$ והשטח בצד ימין של $z_{1-\alpha/2}$ שווה ל $\alpha/2$.

הטווח $[a,b] = [\ \bar{X} - z_{1-\alpha/2}\sigma/\sqrt{n}\ ,\ \bar{X} + z_{1-\alpha/2}\sigma/\sqrt{n}\ ]$ נקרא **רווח סמך** וההסתברות $1-\alpha$ נקרא **רמת מובהקות** .

כדי למצוא את הטווח בשאלה, אנחנו זוכרים כי השטח התחום בגרף של משתנה מקרי כלשהו הוא שווה דווקא להסתברות כי המ"מ נמצא בטווח זו. ז"א השטח התחום של הגרף של $\bar{X}$ בטווח $[a,b]$ הוא שווה ל $P(a \le \bar{X} \le b)$. אבל בשל העובדה ש $\bar{X}$ הוא מתפלג נורמלי (לפי המשפט הגבול המרכזי) אזי ניתן להגדיר משתנה מקרי נורמלי סטנדרדי המתאים, והשטח זו יהיה שווה לשטח התחום של הגרף של Z בטווח המתאים.

נגדיר המשתנה Z להיות

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}\ . \tag{*1}$$

יהי

$$z_{1-\alpha/2}$$

הערך של Z אשר עבורו השטח התחום בטווח $[-z_{1-\alpha/2}, z_{1-\alpha/2}]$ של הגרף [עיין אייור] שווה ל $(1-\alpha)$ והשטח בצד ימין שלו הוא $\alpha/2$ כמתואר באייור להלן.

כמו כן הטווח הנדרש של μ כדי לתת רמת מובהקות $1-\alpha$ (הסתברות $(1-\alpha)$ ל- μ לפול בטווח זו) נמצא ע"י למצוא הטווח המתאים של Z כך שהשטח התחום שווה ל- $(1-\alpha)$. כמו כן לפי הגרף נמצא ש

$$P(-z_{1-\alpha/2} < Z < z_{1-\alpha/2}) = 1-\alpha\ ,$$

על כן

$$P\left(-z_{1-\alpha/2} < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < z_{1-\alpha/2}\right) = 1-\alpha\ .$$

מכפילים אגף הימין ואגף השמאול ב $\sigma/\sqrt{n}$ ולוקחים $\bar{X}$ מכל אגף ואגף ומקבלים כי

$$P\left(\bar{X} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha\ .$$

במילים, לוקחים מדגם מקרי של אורך n מתוך אוכלוסיה בעל שונות σ^2 ידועה. רווח סמך של $100(1-\alpha)\%$ לתוחלת μ נתון ע"י

$$\bar{X} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \tag{*2}$$

כאשר $z_{1-\alpha/2}$ הוא הערך של Z אשר עבורו יש שטח של $\alpha/2$ בצד הימין שלו.
פורמלית:

14.8 חוק. (רווח סמך מדגם מקרי σ^2 ידועה)

אם $\bar{x}$ הוא התוחלת של מדגם מקרי של אורך n מתוך אוכלוסיה בעל שונות σ^2 ידועה, רווח סמך של $100(1-\alpha)\%$ להתוחלת μ נתון ע"י

$$\bar{x} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}$$

כאשר $z_{1-\alpha/2}$ הוא הערך של z אשר עבורו יש שטח של $\alpha/2$ בצד הימין שלו.

14.9 דוגמא. (רווח סמך מדגם מקרי σ^2 ידועה)

הריכוז הממוצע של חמצן ממדגם של מדידות הנלקחות מ 36 מקומות שונים בנהר הוא 2.6 [gr/mm]. מהו הרווח סמך של הרמות המבהוקות של 95% ו 99% להתוחלת של הריכוז חמצן בהנהר בשאלה. יש להניח שהסטיית התקן של האוכלוסיה הוא 0.3 [gr/mm].

פיתרון.
הממוצע של המדגם מקרי הוא
$$\bar{x} = 2.6 \ ,$$
ו-
$$1-\alpha = 0.95 \qquad \Leftrightarrow \qquad \alpha = 0.05 \ .$$
הערך של z אשר עבורו השטח בצד הימין שלו הוא $\alpha/2 = 0.025$, הוא
$$z_{1-\alpha/2} = z_{0.975} = 1.96 \ .$$
מהטבלה. לכן הרווח סמך של 95%, לפי נוסחאה (2*), הוא
$$2.6 - (1.96)\left(\frac{0.3}{\sqrt{36}}\right) < \mu < 2.6 + (1.96)\left(\frac{0.3}{\sqrt{36}}\right) \ .$$
ניתן לצמצם זה ל
$$2.50 < \mu < 2.70 \ .$$
למצוא הרווח סמך של רמת מובהקות של 99%, שים לב ש
$$1-\alpha = 0.99 \qquad \Leftrightarrow \qquad \alpha = 0.01, \qquad \Leftrightarrow \qquad \alpha/2 = 0.005 \ , \qquad \Leftrightarrow \qquad 1-\alpha/2 = 0.995 \ .$$
יש לחפש את הערך של z כך שבצד הימין שלו יש שטח של $\alpha/2 = 0.005$ מהטבלה הערך הנדרש הוא $z_{1-\alpha/2} = z_{0.995} = 2.575$ ולכן הרווח סמך של רמת מובהקות של 99% הוא
$$2.6 - (2.575)\left(\frac{0.3}{\sqrt{36}}\right) < \mu < 2.6 + (2.575)\left(\frac{0.3}{\sqrt{36}}\right) \ ,$$
או
$$2.47 < \mu < 2.73 \ .$$
שימו לב יש צורך לרווח יותר ארוך כדי להשיג ערך של μ יותר מדוייק. ■

הרווח סמך נותן הדייוק של האומדן של μ. אם μ נמצא במרכז של הרווח, אז $\bar{x}$ מעריך את μ ללא שגיאה. רוב הזמן אבל, $\bar{x}$ לא יהיה שווה בדיוק ל μ, כך שיהיה שגיאה בין האומדן לבין הערך המדויק של μ. הגודל של השגיאה זו הוא שווה להערך מוחלט של ההפרש בין μ לבין $\bar{x}$, וניתן להיות $100(1-\alpha)$% בטוח כי ההפרש זו לא יעבור $z_{1-\alpha/2}\sigma/\sqrt{n}$. אפשר לראות את זה עם העזרה של האייור להלן.

14.10 מסקנה. (סמך באומדן של μ)
אם לוקחים $\bar{x}$ להיות אומדן של μ, אז יש רמת מובהקות של $100(1-\alpha)$% שהשגיאה לא יעבור $z_{1-\alpha/2}\sigma/\sqrt{n}$.

בדוגמה לעייל יש רמת מובהקות של 95% שההפרש בין התוחלת של המדגם $\bar{x} = 2.6$ ו μ לא יעבור $(1.96)(0.3)/\sqrt{36} = 0.13$. לעתים יש צורך לדעת את האורך הנדרש של המדגם כדי לוודע שהשגיאה באומדן של μ לא יעבור ערך נתון e. על ידי המסקנה 14.10 לעייל, יש צורך לבחור n כך ש
$$z_{1-\alpha/2}\sigma/\sqrt{n} = e \ .$$
פותרים את המשוואה זו כדי לקבל נוסחאה ל n:

14.11 מסקנה. (סמך באומדן של μ)
אם לוקחים $\bar{x}$ להיות אומדן של μ, אז יש רמת מובהקות של $100(1-\alpha)\%$ שהשגיאה לא יעבור e כאשר האורך של המדגם הוא

$$n = \left(\frac{z_{1-\alpha/2}\sigma}{e}\right)^2 \ .$$

14.12 דוגמא. מהו האורך הנדרש של המדגם להשיג רמת מובהקות של 95% שיש לאומדן של μ בדוגמה 14.9 שגיאה פחות מ 0.05?

פיתרון. הסטיית התקן הוא $\sigma = 0.3$. לכן לפי מסקנה 14.11:

$$n = \left(\frac{(1.96)(0.3)}{(0.05)}\right)^2 = 138.3 \ .$$

לכן אפשר להיות 95% בטוח שמדגם מקרי של אורך של $n = 139$ יתן אומדן $\bar{x}$ של μ עם שגיאה פחות מ 0.05. ■

טבלות של ערכים של התפלגויות

$\Phi(z)$	z
0.5000000	0.0000000
0.5500000	0.1256613
0.6000000	0.2533471
0.6500000	0.3853205
0.7000000	0.5244005
0.7500000	0.6744898
0.8000000	0.8416212
0.8500000	1.0364334
0.9000000	1.2815516
0.9100000	1.3407550
0.9200000	1.4050716
0.9300000	1.4757910
0.9400000	1.5547736
0.9500000	1.6448536
0.9600000	1.7506861
0.9700000	1.8807936
0.9800000	2.0537489
0.9900000	2.3263479
0.9950000	2.5758293
0.9990000	3.0902323
0.9995000	3.2905267
0.9999000	3.7190165
0.9999500	3.8905919
0.9999900	4.2648908
0.9999950	4.4171734
0.9999990	4.7534243
0.9999995	4.8916385
0.9999999	5.1993376

n	$t_{0.995}$	$t_{0.990}$	$t_{0.975}$	$t_{0.950}$	$t_{0.900}$	$t_{0.800}$	$t_{0.750}$	$t_{0.700}$	$t_{0.600}$	$t_{0.550}$
1.000	63.657	31.821	12.706	6.314	3.078	1.376	1.000	0.727	0.325	0.158
2.000	9.925	6.965	4.303	2.920	1.886	1.061	0.816	0.617	0.289	0.142
3.000	5.841	4.541	3.182	2.353	1.638	0.978	0.765	0.584	0.277	0.137
4.000	4.604	3.747	2.776	2.132	1.533	0.941	0.741	0.569	0.271	0.134
5.000	4.032	3.365	2.571	2.015	1.476	0.920	0.727	0.559	0.267	0.132
6.000	3.707	3.143	2.447	1.943	1.440	0.906	0.718	0.553	0.265	0.131
7.000	3.499	2.998	2.365	1.895	1.415	0.896	0.711	0.549	0.263	0.130
8.000	3.355	2.896	2.306	1.860	1.397	0.889	0.706	0.546	0.262	0.130
9.000	3.250	2.821	2.262	1.833	1.383	0.883	0.703	0.543	0.261	0.129
10.000	3.169	2.764	2.228	1.812	1.372	0.879	0.700	0.542	0.260	0.129
11.000	3.106	2.718	2.201	1.796	1.363	0.876	0.697	0.540	0.260	0.129
12.000	3.055	2.681	2.179	1.782	1.356	0.873	0.695	0.539	0.259	0.128
13.000	3.012	2.650	2.160	1.771	1.350	0.870	0.694	0.538	0.259	0.128
14.000	2.977	2.624	2.145	1.761	1.345	0.868	0.692	0.537	0.258	0.128
15.000	2.947	2.602	2.131	1.753	1.341	0.866	0.691	0.536	0.258	0.128
16.000	2.921	2.583	2.120	1.746	1.337	0.865	0.690	0.535	0.258	0.128
17.000	2.898	2.567	2.110	1.740	1.333	0.863	0.689	0.534	0.257	0.128
18.000	2.878	2.552	2.101	1.734	1.330	0.862	0.688	0.534	0.257	0.127
19.000	2.861	2.539	2.093	1.729	1.328	0.861	0.688	0.533	0.257	0.127
20.000	2.845	2.528	2.086	1.725	1.325	0.860	0.687	0.533	0.257	0.127
21.000	2.831	2.518	2.080	1.721	1.323	0.859	0.686	0.532	0.257	0.127
22.000	2.819	2.508	2.074	1.717	1.321	0.858	0.686	0.532	0.256	0.127
23.000	2.807	2.500	2.069	1.714	1.319	0.858	0.685	0.532	0.256	0.127
24.000	2.797	2.492	2.064	1.711	1.318	0.857	0.685	0.531	0.256	0.127
25.000	2.787	2.485	2.060	1.708	1.316	0.856	0.684	0.531	0.256	0.127
26.000	2.779	2.479	2.056	1.706	1.315	0.856	0.684	0.531	0.256	0.127
27.000	2.771	2.473	2.052	1.703	1.314	0.855	0.684	0.531	0.256	0.127
28.000	2.763	2.467	2.048	1.701	1.313	0.855	0.683	0.530	0.256	0.127
29.000	2.756	2.462	2.045	1.699	1.311	0.854	0.683	0.530	0.256	0.127
30.000	2.750	2.457	2.042	1.697	1.310	0.854	0.683	0.530	0.256	0.127
40.000	2.704	2.423	2.021	1.684	1.303	0.851	0.681	0.529	0.255	0.126
60.000	2.660	2.390	2.000	1.671	1.296	0.848	0.679	0.527	0.254	0.126
120.000	2.617	2.358	1.980	1.658	1.289	0.845	0.677	0.526	0.254	0.126

עשרה: הוכחה של המשפט הגבול מרכזי

14.13 חוק. (משפט הגבול המרכזי) יהי X משתנה מקרי בעל תוחלת μ וסטיית התקן σ, ויהי

$$x_1, \ldots, x_n$$

מדגם מקרי מתוך X. אזי,כאשר $n \geq 30$ או לכל n במקרה ש X מתפלג נורמלית, מתקיים

$$\sum_{i=1}^{n} X_i \sim N\left(n\mu, n\sigma^2\right),$$

או במילים אחרות, כאשר $\bar{X} := \frac{1}{n}\sum_{i=1}^{n} X_i$ הוא התוחלת של המדגם,

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) .$$

14.14 מסקנה. (משפט הגבול המרכזי) יהי $\bar{X}$ התוחלת של מדגם מקרי של אורך n מתוך אוכלוסיה בעל תוחלת μ ושונות σ^2. אזי,כאשר $n \geq 30$, ההתפלגות של המשתנה

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

היא ההתפלגות הנורמלית הסטנדרדית $n(z, 0, 1)$, קרי

$$Z \sim N(0, 1) .$$

הוכחה.

The central limit theorem (CLT) states that when independent random variables are added, their properly normalized sum tends toward a normal distribution (a bell curve) even if the original variables themselves are not normally distributed.

Let $\{X_1, \ldots, X_n\}$ be a random sample of size n, that is, a sequence of independent and identically distributed (IID)[1] random variables drawn from a distribution of expected value μ and finite variance σ^2. The sample average

$$\bar{X}_n := \frac{X_1 + \cdots + X_n}{n},$$

by the law of large numbers[2] converges to the expected value μ as $n \to \infty$.

Assume $\{X_1, \ldots, X_n\}$ are independent and identically distributed random variables, each with mean μ abd finite variance σ^2. The sum $X_1 + \cdots + X_n$ has mean $n\mu$ and variance $n\sigma^2$. Define the random variable

$$Z_n := \frac{X_1 + \cdots + X_n - n\mu}{\sqrt{n\sigma^2}} = \sum_{i=1}^{n}\left(\frac{X_i - \mu}{\sqrt{n\sigma^2}}\right) = \sum_{i=1}^{n} \frac{1}{\sqrt{n}} Y_i, \qquad \text{(\#1)}$$

[1] A set of random variables is independent and identically distributed if each random variable has the same probability distribution as the others and all are mutually independent.

[2] The law of large numbers (LLN) posits that the average of the results obtained from a large number of trials should be close to the expected value, and will tend to become closer to the expected value as more trials are performed

where $Y_i := (X_i - \mu)/\sigma$, which has zero mean and unit variance: $\mathrm{var}(Y) = E\left[(Y_i - \text{mean of } Y)^2\right] = E\left[(Y_i)^2\right] = \frac{1}{\sigma^2} E\left[(X_i - \mu)^2\right] = \sigma^2/\sigma^2 = 1$. The characteristic function of Z_n is given by

$$\varphi(t)_{Z_n} = \varphi(t)_{\sum_{j=1}^n n^{-1/2} Y_j} = E\left[\exp\left(it \sum_{j=1}^n n^{-1/2} Y_j\right)\right] = E\left[\prod_{j=1}^n \exp\left(itn^{-1/2} Y_j\right)\right]. \quad \text{(\#2)}$$

By assumption the Y_j are identically distributed, which means they each have the same expectation value, and it follows from (#2) that

$$\varphi(t)_{Z_n} = \prod_{j=1}^n E\left[\exp\left(itn^{-1/2} Y_j\right)\right] = \varphi\left(n^{-1/2} t\right)_{Y_1} \varphi\left(n^{-1/2} t\right)_{Y_2} \cdots \varphi\left(n^{-1/2} t\right)_{Y_n} = \left[\varphi\left(n^{-1/2} t\right)_{Y_1}\right]^n. \quad \text{(\#3)}$$

The characteristic function of Y_1 is, by Taylor's theorem, $E\left[e^{itn^{-1/2} Y_1}\right] = E[1] + E\left[in^{-1/2} Y_1\right] t + E\left[-\left(n^{-1/2} Y_1\right)^2\right] t^2/2 + O\left(t^3\right)$. Above it was established that Y_1 has zero expectation value ($E[Y_1] = 0$) and variance one ($E\left[(Y_1)^2\right] = 1$), which means that

$$\varphi\left(n^{-1/2} t\right)_{Y_1} = E\left[e^{itn^{-1/2} Y_1}\right] = 1 - \frac{t^2}{2n} + \frac{t^4}{4! n^2} + o\left(\frac{t^4}{n^2}\right), \quad \text{(\#4)}$$

where $o\left(t^4/n^2\right)$ means something that goes to zero more rapidly than t^4/n^2. Hence,

$$\begin{aligned} \varphi(t)_{Z_n} &= \left(\varphi\left(n^{-1/2} t\right)_{Y_1}\right)^n \\ &= \left(1 - \frac{t^2}{2n}\right)^n + n\left(1 - \frac{t^2}{2n}\right)^{n-1} \left(\frac{t^4}{4! n^2} + o\left(\frac{t^4}{n^2}\right)\right) + \cdots \end{aligned} .$$

The form of the exponential function as a limit is $e^x = \lim_{n\to\infty} (1 + x/n)^n$. It follows that

$$\lim_{n\to\infty} \varphi(t)_{Z_n} = e^{-t^2/2} \lim_{n\to\infty} \left(1 + n\left(1 - \frac{t^2}{2n}\right)^{-1} \left(\frac{t^4}{4! n^2} + o\left(\frac{t^4}{n^2}\right)\right) + \cdots\right) .$$

All of the higher order terms inside the brackets vanish in the limit $n \to \infty$. Therefore,

$$\lim_{n\to\infty} \varphi(t)_{Z_n} = e^{-t^2/2}. \quad \text{(\#5)}$$

The right hand side equals the characteristic function of a standard normal distribution $N(0,1)$ **(prove)**, which implies through Le□vy's continuity theorem **(state and prove)** that the distribution of Z_n will approach $N(0,1)$ as $n \to \infty$. Therefore, the sample average

$$\bar{X}_n = \frac{X_1 + \cdots + X_n}{n}$$

is such that

$$\frac{\sqrt{n}}{\sigma}\left(\bar{X}_n - \mu\right)$$

converges to the normal distribution $N(0,1)$, from which the central limit theorem follows.

■