Exercises in Introduction to Mathematical Statistics (Ch. 5)

Tomoki Okuno

September 14, 2022

Note

- Not all solutions are provided: Exercises that are too simple or not very important to me are skipped.
- Texts in red are just attentions to me. Please ignore them.

5 Consistency and Limiting Distributions

5.1 Convergence in Probability

5.1.1. Let a_n be a sequence of real numbers. Hence, we can also say that a_n is a sequence of constant (degenerate) random variables. Let a be a real number. Show that $a_n \to a$ is equivalent to $a_n \xrightarrow{P} a$.

Solution.

$$\begin{aligned} a_n \to a &\Leftrightarrow \forall\epsilon > 0, \exists N \in \mathbb{N} \text{ such that } |a_n - a| \le \epsilon \text{ for } n > N \\ &\Leftrightarrow P(|a_n - a| \le \epsilon) \to 1 \text{ as } n \to \infty \\ &\Leftrightarrow P(|a_n - a| > \epsilon) \to 0 \text{ as } n \to \infty \\ &\Leftrightarrow a_n \xrightarrow{P} a. \end{aligned}$$

5.1.2. Let the random variable Y_n have a distribution that is $Binomial(n, p)$.

(a) Prove that Y_n/n converges in probability to p. This result is one form of the weak law of large numbers.

Solution.
Let $X_1, ..., X_n \overset{iid}{\sim}$ Bernoulli(p) with $\mu = p$ and $\sigma^2 = p(1-p)$. Since $Y_n = \sum_{i=1}^n X_i$,

$$Y_n/n = \sum_{i=1}^n X_i/n = \overline{X}_n \xrightarrow{P} p \quad \text{by WLLN.}$$

(b) Prove that $1 - Y_n/n$ converges in probability to $1 - p$.

Solution. Let $g(x) = 1 - x$, which is continuous at all x. Then $1 - Y_n/n = g(Y_n/n) \xrightarrow{P} g(p) = 1 - p$.

(c) Prove that $Y_n/n(1 - Y_n/n)$ converges in probability to $p(1 - p)$.

Solution. Let $g(x) = x(1 - x)$. Then, $Y_n/n(1 - Y_n/n) = g(Y_n/n) \xrightarrow{P} g(p) = p(1 - p)$.

5.1.3. Let W_n denote a random variable with mean μ and variance b/n^p, where $p > 0$, μ, and b are constants (not functions of n). Prove that W_n converges in probability to μ.

Solution.

By Chebyshev's inequality, for $\forall \epsilon > 0$,

$$P(|W_n - \mu| \geq \epsilon) \leq \frac{\text{Var}(W_n)}{\epsilon^2} = \frac{b}{\epsilon^2 n^p} \to 0 \quad \text{as } n \to \infty.$$

5.1.4. Derive the cdf given in expression (5.1.1).

Solution.

The cdf of X is $F_X(x) = 0,\ x \leq 0;\ x/\theta,\ 0 < x \leq \theta;\ 1,\ x > \theta$. Hence,

$$F_{Y_n}(t) = P(Y_n < t) = P(X_i < t,\ i = 1, .., n) = [F(t)]^n = \begin{cases} 0 & t \leq 0 \\ \left(\frac{t}{\theta}\right)^n & 0 < t \leq \theta \\ 1 & t > \theta. \end{cases}$$

5.1.7. Let $X_1, ..., X_n$ be iid random variables with common pdf

$$f(x) = \begin{cases} e^{-(x-\theta)} & x > \theta,\ -\infty < \theta < \infty \\ 0 & \text{elsewhere.} \end{cases}$$

This pdf is called the **shifted exponential**. Let $Y_n = \min\{X_1, ..., X_n\}$. Prove that $Y_n \to \theta$ in probability by first obtaining the cdf of Y_n.

Solution.

$$\begin{aligned} P(Y_n \geq y) &= P(\min\{X_1, ..., X_n\} \geq y) \\ &= P(X_i \geq y,\ i = 1, \ldots, n) \\ &= [P(X \geq y)]^n \quad \text{since } X_1, ..., X_n \text{ are iid} \\ &= \begin{cases} 1 & y \leq \theta \\ \left[\int_y^\infty e^{-(t-\theta)} dt\right]^n = e^{-n(y-\theta)} & y > \theta. \end{cases} \end{aligned}$$

Hence, the cdf of Y_n is

$$F_{Y_n}(y) = 1 - P(Y_n \geq y) = \begin{cases} 0 & y \leq \theta \\ 1 - e^{-n(y-\theta)} & y > \theta. \end{cases}$$

Let $\epsilon > 0$ be given.

$$\begin{aligned} P(|Y_n - \theta| \leq \epsilon) &= P(-\epsilon < Y_n - \theta < \epsilon) \\ &= P(\theta - \epsilon < Y_n < \theta + \epsilon) \\ &= F(\theta + \epsilon) - F(\theta - \epsilon) \\ &= (1 - e^{-n\epsilon}) - 0 \\ &\to 1 \quad \text{as } n \to \infty, \end{aligned}$$

which means $Y_n \xrightarrow{P} \theta$.

5.1.8. Using the assumptions behind the confidence interval given in expression (4.2.9), show that

$$\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} \Big/ \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \xrightarrow{P} 1$$

Solution.

Let

$$g(x,y) = \sqrt{\frac{x}{n_1} + \frac{y}{n_2}} \Big/ \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}.$$

Since $S_1^2 \xrightarrow{P} \sigma_1^2$ and $S_2^2 \xrightarrow{P} \sigma_2^2$ and $g(x,y)$ is continuous at all (x,y),

$$\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} \Big/ \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = g(S_1^2, S_2^2) \xrightarrow{P} g(\sigma_1^2, \sigma_2^2) = 1.$$

5.1.9. For Exercise 5.1.7, obtain the mean of Y_n. Is Y_n an unbiased estimator of θ? Obtain an unbiased estimator of θ based on Y_n.

Solution.

First, we obtain the pdf, $f_{Y_n}(y)$:

$$f_{Y_n}(y) = F'_{Y_n}(y) = \begin{cases} 0 & y \le \theta \\ ne^{-n(y-\theta)} & y > \theta. \end{cases}$$

Hence, the mean of Y_n is

$$\begin{aligned} E(Y_n) &= n\int_\theta^\infty ye^{-n(y-\theta)}dy \\ &= \int_0^\infty \left(\frac{t}{n} + \theta\right) e^{-t}dt \quad (t = n(y-\theta)) \\ &= \frac{1}{n} + \theta \quad \text{since} \int_0^\infty te^{-t}dt = \int_0^\infty e^{-t}dt = 1. \end{aligned}$$

Hence, Y_n is biased for θ, but $Y_n - 1/n$ is unbiased since $E(Y_n - 1/n) = E(Y_n) - 1/n = \theta$.

5.2 Convergence in Distribution

Simple example that illustrates the difference between two convergences

Let X be a continuous random variable with a pdf $f_X(x)$ that is symmetric at 0. Then it is easy to show that the pdf of $-X$ is also $f_X(x)$. Thus, X and $-X$ have the same distributions. Define the sequence of random variables X_n as

$$X_n = \begin{cases} X & \text{if } n \text{ is odd} \\ -X & \text{if } n \text{ is even}. \end{cases}$$

Clearly, $F_{X_n} = F_X$ for all x in the support of X, so that $X_n \xrightarrow{D} X$. On the other hand, X does not get close to X. In particular, $X_n \overset{P}{\nrightarrow} X$ in probability.

5.2.1. Let $\overline{X}_n$ denote the mean of a random sample of size n from a distribution that is $N(\mu, \sigma^2)$. Find the limiting distribution of $\overline{X}_n$.

Solution. Since $\overline{X}_n \sim N(\mu, \sigma^2/n)$, $\text{Var}(\overline{X}_n) \to 0$ as $n \to \infty$, indicating that $\overline{X}_n$ is generate at μ.

5.2.2. Let Y_1 denote the minimum of a random sample of size n from a distribution that has pdf $f(x) = e^{-(x-\theta)}, \theta < x < \infty$, zero elsewhere. Let $Z_n = n(Y_1 - \theta)$. Investigate the limiting distribution of Z_n.

Solution.

By Exercise 5.1.7,

$$F_{Y_1}(y) = \begin{cases} 0 & y \le \theta \\ 1 - e^{-n(y-\theta)} & y > \theta. \end{cases}$$

Thus,

$$\begin{aligned} F_{Z_n}(z) = P(Z \le z) &= P(n(Y_1-\theta) \le z) \\ &= P(Y_1 \le z/n + \theta) \\ &= \begin{cases} 0 & z \le 0 \\ 1 - e^{-z} & z > 0 \end{cases} \end{aligned}$$

which holds for $\forall n$. Then $f_{Z_n}(z) = e^{-z}$, or $Z_n \xrightarrow{D} \Gamma(1,1)$.

5.2.3. Let Y_n denote the maximum of a random sample of size n from a distribution of the continuous type that has cdf $F(x)$ and pdf $f(x) = F'(x)$. Find the limiting distribution of $Z_n = n[1 - F(Y_n)]$.

Solution.

$$F_{Y_n}(y) = P(Y_n \le y) = [P(X \le y)]^n = [F_X(y)]^n.$$

Hence

$$\begin{aligned} F_{Z_n}(z) &= P(n[1-F(Y_n)] \le z) \\ &= P[F(Y_n) \ge 1 - z/n] \\ &= P[Y_n \ge F_X^{-1}(1-z/n)] \quad \text{since } Y_n \text{ is nondecreasing} \\ &= 1 - F_{Y_n}\{F_X^{-1}(1-z/n)\} \\ &= 1 - [F_X\{F_X^{-1}(1-z/n)\}]^n \\ &= 1 - (1-z/n)^n \to 1 - e^{-z} \text{ as } n \to \infty, \end{aligned}$$

which means that $Z_n \xrightarrow{D} \Gamma(1,1)$.

5.2.4. Let Y_2 denote the second smallest item of a random sample of size n from a distribution of the continuous type that has cdf $F(x)$ and pdf $f(x) = F'(x)$. Find the limiting distribution of $W_n = nF(Y_2)$.

Solution.

The pdf of Y_2 is

$$f_{Y_2}(y) = n(n-1)F_X(y)[1-F_X(y)]^{n-2}f_X(y).$$

Hence

$$\begin{aligned} F_{W_n}(w) &= P(F(Y_2) \le w/n) \\ &= P(Y_2 \le F^{-1}(w/n)) \\ &= \int_{-\infty}^{F^{-1}(w/n)} n(n-1)F_X(y)[1-F_X(y)]^{n-2}f_X(y)dy \\ &= -nF_X(y)[1-F_X(y)]^{n-1}\Big|_{-\infty}^{F^{-1}(w/n)} + \int_{-\infty}^{F^{-1}(w/n)} nf_X(y)[1-F_X(y)]^{n-1}dy \\ &= -w[1-w/n]^{n-1} - [1-F_X(y)]^n\Big|_{-\infty}^{F^{-1}(w/n)} \\ &= -w[1-w/n]^{n-1} + [1-(1-w/n)^n] \\ &\to -we^{-w} + (1-e^{-w}) \equiv F_W(w), \\ f_W(w) = F_W'(w) &= -e^{-w} + we^{-w} + e^{-w} = we^{-w}, \end{aligned}$$

which means $W_n \xrightarrow{D} W \sim \Gamma(2,1)$.

5.2.5. Let the pmf of Y_n be $p_n(y) = 1$, $y = n$, zero elsewhere. Show that Y_n does not have a limiting distribution. (In this case, the probability has "escaped" to infinity.)

Solution.

The cdf of Y_n is

$$F_{Y_n}(y) = \begin{cases} 0 & y < n \\ 1 & y \geq n. \end{cases}$$

Assume Y is the limiting distribution of Y_n,

$$F_Y(y) = \lim_{n\to\infty} F_{Y_n}(y) = 0, \quad -\infty < y < \infty.$$

Since $F_Y(y) \neq F_{Y_n}(y)$, Y does not exist.

5.2.6. Let $X_1, X_2, ..., X_n$ be a random sample of size n from a distribution that is $N(\mu, \sigma^2)$, where $\sigma^2 > 0$. Show that the sum $Z_n = \sum_1^n X_i$ does not have a limiting distribution.

Solution.

Since $X_1, ..., X_n$ are iid $N(\mu, \sigma^2)$,

$$Z_n = \sum_i X_i \sim N(n\mu, n\sigma^2).$$

Thus, the cdf of Z_n is

$$\begin{aligned} F_{Z_n}(z) = P(Z_n \leq z) &= P\left(\frac{Z_n - n\mu}{\sqrt{n\sigma^2}} \leq \frac{z - n\mu}{\sqrt{n\sigma^2}}\right) \\ &= \Phi\left(\frac{z - n\mu}{\sqrt{n\sigma^2}}\right) \\ &\to \begin{cases} \Phi(-\infty) = 0 & \mu > 0 \\ \Phi(0) = 1/2 & \mu = 0 \\ \Phi(\infty) = 1 & \mu < 0 \end{cases} \text{ as } n \to \infty, \end{aligned}$$

which means that Z_n does not have a limiting distribution.

5.2.7. Let X_n have a gamma distribution with parameter $\alpha = n$ and β, where β is not a function of n. Let $Y_n = X_n/n$. Find the limiting distribution of Y_n.

Solution.

$$M_{X_n}(t) = (1 - \beta t)^{-n} \Rightarrow M_{Y_n}(t) = M_{X_n}(t/n) = (1 - \beta t/n)^{-n} \to e^{\beta t} \quad \text{as } n \to \infty,$$

which indicates that the limiting distribution of Y_n is degenerate at β.

5.2.8. Let Z_n be $\chi^2(n)$ and let $W_n = Z_n/n^2$. Find the limiting distribution of W_n.

Solution.

Since $M_{Z_n}(t) = (1 - 2t)^{-n/2}$,

$$\begin{aligned} M_{W_n}(t) = M_{Z_n}(t/n^2) &= (1 - 2t/n^2)^{-n/2} \\ &= \left(1 + \frac{\sqrt{2t}}{n}\right)^{-n/2} \left(1 - \frac{\sqrt{2t}}{n}\right)^{-n/2} \\ &\to \exp\left(-\frac{\sqrt{2t}}{2}\right) \exp\left(\frac{\sqrt{2t}}{2}\right) \quad \text{as } n \to \infty \\ &= e^0, \end{aligned}$$

which implies that the limiting distribution of W_n is degenerate at 0.

5.2.11. Let $p = 0.95$ be the probability that a man, in a certain age group, lives at least 5 years.

(a) If we are to observe 60 such men and if we assume independence, use R to compute the probability that at least 56 of them live 5 or more years.

Solution.

Let $X \sim \text{Binomial}(60, 0.95)$, $P(X \geq 56) = \texttt{1 - pbinom(55, 60, 0.95)} = 0.820$.

(b) Find an approximation to the result of part (a) by using the Poisson distribution.

Solution.

Let $Y = 60 - X \sim \text{Binomial}(60, 0.05)$, then $Y \overset{D}{\sim} \text{Poisson}(3)$. Using this approximation to obtain

$$P(X \geq 56) = P(Y \leq 4) = \texttt{ppois(4, 3)} = 0.815,$$

which is close to the exact probability.

5.2.12. Let the random variable Z_n have a Poisson distribution with parameter $\mu = n$. Show that the limiting distribution of the random variable $Y_n = (Z_n - n)/\sqrt{n}$ is normal with mean zero and variance 1.

Solution.

$$\begin{aligned}
M_{Z_n}(t) &= e^{n(e^t - 1)}, \\
M_{Y_n}(t) &= e^{-\sqrt{n}t} M_{Z_n}(t/\sqrt{n}) \\
&= e^{-\sqrt{n}t} e^{n(e^{t/\sqrt{n}} - 1)} \\
&= e^{-\sqrt{n}t} e^{n((1 + t/\sqrt{n} + t^2/2n + O(t^3/n\sqrt{n}) - 1)} \\
&\to e^{t^2/2} \quad \text{as } n \to \infty,
\end{aligned}$$

indicating that $Y_n \xrightarrow{D} N(0,1)$. This is a specific case of the Central Limit Theorem.

5.2.15. Let $\overline{X}_n$ denote the mean of a random sample of size n from a Poisson distribution with parameter $\mu = 1$.

(a) Show that the mgf of $Y_n = \sqrt{n}(X_n - \mu)/\sigma = \sqrt{n}(X_n - 1)$ is given by $\exp[-t\sqrt{n} + n(e^{t/\sqrt{n}} - 1)]$.

Solution.

The mgf of $X_i \overset{iid}{\sim} \text{Poisson}(1)$ is $M_X(t) = e^{e^t - 1} \Rightarrow M_{\bar{X}}(t) = [M_X(t/n)]^n = e^{n(e^{t/n} - 1)}$. Hence, $M_Y(t) = e^{-\sqrt{n}t} M_{\bar{X}}(\sqrt{n}t) = e^{-\sqrt{n}t} e^{n(e^{t/\sqrt{n}} - 1)} = \exp[-t\sqrt{n} + n(e^{t/\sqrt{n}} - 1)]$.

(b) Investigate the limiting distribution of Y_n as $n \to \infty$.

Solution. By CLT, $\sqrt{n}(\overline{X}_n - 1) \xrightarrow{D} N(0,1)$.

5.2.16. Using Exercise 5.2.15 and the Δ-method, find the limiting distribution of $\sqrt{n}(\sqrt{\overline{X}_n} - 1)$.

Solution.

Let $g(x) = \sqrt{x}$, which is continuous at all x. Then, by the Delta method,

$$\sqrt{n}(g(\overline{X}_n) - g(1)) \xrightarrow{D} N(0, \{g'(1)\}^2) \;\Rightarrow\; \sqrt{n}(\sqrt{\overline{X}_n} - 1) \xrightarrow{D} N(0, 1/4).$$

5.2.17. Let $\overline{X}_n$ denote the mean of a random sample of size n from a distribution that has pdf $f(x) = e^{-x}$, $0 < x < \infty$, zero elsewhere.

(a) Show that the mgf $M_{Y_n}(t)$ of $Y_n = \sqrt{n}(\overline{X}_n - 1)$ is

$$M_{Y_n}(t) = [e^{t/\sqrt{n}} - (t/\sqrt{n})e^{t/\sqrt{n}}]^{-n}, \quad t < \sqrt{n}.$$

Solution.

Since $M_X(t) = (1-t)^{-1}$, $t < 1$,

$$M_{\overline{X}_n}(t) = [M_X(t/n)]^n = (1-t/n)^{-n}.$$

Hence,

$$M_{Y_n}(t) = e^{-t\sqrt{n}} M_{\overline{X}_n}(\sqrt{n}t) = e^{-t\sqrt{n}}(1-t/\sqrt{n})^{-n} = [e^{t/\sqrt{n}} - (t/\sqrt{n})e^{t/\sqrt{n}}]^{-n}.$$

(b) Find the limiting distribution of Y_n as $n \to \infty$.

Solution.

By CLT, $Y_n \xrightarrow{D} N(0,1)$. Another solution is:

$$\begin{aligned}
M_{Y_n}(t) &= [e^{t/\sqrt{n}} - (t/\sqrt{n})e^{t/\sqrt{n}}]^{-n} \\
&= \left[e^{t/\sqrt{n}}\left(1 - \frac{t}{\sqrt{n}}\right)\right]^{-n} \\
&\approx \left[\left(1 + \frac{t}{\sqrt{n}} + \frac{t^2}{2n}\right)\left(1 - \frac{t}{\sqrt{n}}\right)\right]^{-n} \\
&= \left[1 - \frac{t^2}{n} + \frac{t^2}{2n} - \frac{t^3}{2n\sqrt{n}}\right]^{-n} \\
&= \left[1 - \frac{t^2}{2n} - \frac{t^3}{2n\sqrt{n}}\right]^{-n} \\
&\to e^{(-t^2/2)(-1)} = e^{t^2/2}.
\end{aligned}$$

5.2.18. Continuing with Exercise 5.2.17, use the Δ-method to find the limiting distribution of $\sqrt{n}(\sqrt{\overline{X}_n} - 1)$.

Solution. Exactly the same as 5.2.16.

5.2.19. Let $Y_1 < Y_2 < \cdots < Y_n$ be the order statistics of a random sample (see Section 5.2) from a distribution with pdf $f(x) = e^{-x}$, $0 < x < \infty$, zero elsewhere. Determine the limiting distribution of $Z_n = (Y_n - \log n)$.

Solution.

$$\begin{aligned}
F_{Z_n}(z) &= P(Y_n - \log n \le z) \\
&= P[Y_n \le z + \log n] \\
&= P[X \le z + \log n]^n \\
&= [F_X(z + \log n)]^n \\
&= [1 - e^{-z-\log n}]^n \\
&= [1 - e^{-z}/n]^n \to e^{-e^{-z}}
\end{aligned}$$

as $n \to \infty$.

5.2.20. Let $Y_1 < Y_2 < \cdots < Y_n$ be the order statistics of a random sample (see Section 5.2) from a distribution with pdf $f(x) = 5x^4$, $0 < x < 1$, zero elsewhere. Find p so that $Z_n = n^p Y_1$ converges in distribution.

Solution.

$$\begin{aligned}
F_{Z_n}(z) &= P(n^p Y_1 \le z) = P[Y_1 \le z/n^p] = 1 - P[Y_1 > z/n^p] \\
&= 1 - [1 - P(X \le z/n^p)]^n = 1 - [1 - F(z/n^p)]^n \\
&= 1 - [1 - z^5/n^{5p}]^n \to e^{-z^5} = F_Z(z) \text{ as } n \to \infty,
\end{aligned}$$

which holds only when $5p = 1$ or $p = 1/5$.

5.3 Central Limit Theorem

5.3.1. Let $\overline{X}$ denote the mean of a random sample of size 100 from a distribution that is $\chi^2(50)$. Compute an approximate value of $P(49 < \overline{X} < 51)$.

Solution.

Since $E(X) = 50, \text{Var}(X) = 100$, by CLT, $\sqrt{100}(\overline{X} - 50)/10 = \overline{X} - 50 \xrightarrow{D} N(0,1)$. Hence,

$$P(49 < \overline{X} < 51) = P(-1 < \overline{X} - 50 < 1) = \Phi(1) - \Phi(-1) = 0.683.$$

5.3.2. Let $\overline{X}$ denote the mean of a random sample of size 128 from a gamma distribution with $\alpha = 2$ and $\beta = 4$. Approximate $P(7 < \overline{X} < 9)$.

Solution.

Since $E(X) = \alpha\beta = 8, \text{Var}(X) = \alpha\beta^2 = 32,\ \sqrt{128}(\overline{X} - 8)/\sqrt{32} = 2(\overline{X} - 8) \xrightarrow{D} N(0,1)$ by CLT. Hence,

$$P(7 < \overline{X} < 9) = P(-2 < 2(\overline{X} - 8) - 50 < 2) = \Phi(2) - \Phi(-2) = 0.954.$$

5.3.3. Let Y be $b(72, \frac{1}{3})$. Approximate $P(22 \le Y \le 28)$.

Solution.

Since $E(Y) = np = 24, \text{Var}(X) = np(1-p) = 16$, $(Y - 24)/4 \xrightarrow{D} N(0,1)$ by CLT. Hence,

$$P(22 \le Y \le 28) = P(-0.5 < (Y - 24)/4 < 1) = \Phi(1) - \Phi(-0.5) = 0.532.$$

Note that the official answer uses a continuity correction. Then

$$P(21.5 \le Y \le 28.5) = P(-0.625 < (Y - 24)/4 < 1.125) = \Phi(1.125) - \Phi(-0.625) = 0.604.$$

5.3.4. Compute an approximate probability that the mean of a random sample of size 15 from a distribution having pdf $f(x) = 3x^2$, $0 < x < 1$, zero elsewhere, is between $\frac{3}{5}$ and $\frac{4}{5}$.

Solution.

Since $E(X) = 3/4$ and $E(X^2) = 3/5 \Rightarrow \text{Var}(X) = 3/80$. Hence,

$$\sqrt{n}(\overline{X} - \mu)/\sigma = \frac{\sqrt{15}(\overline{X} - 3/4)}{\sqrt{3/80}} = 20(\overline{X} - 0.75) \xrightarrow{D} N(0,1) \quad \text{by CLT.}$$

Thus,

$$P(3/5 < \overline{X} < 4/5) = P(-3 < 20(\overline{X} - 0.75) < 1) = \Phi(1) - \Phi(-3) = 0.840.$$

5.3.6. Let Y be $b(400, \frac{1}{5})$. Compute an approximate value of $P(0.25 < Y/400)$.

Solution.

$$\frac{Y - np}{\sqrt{np(1-p)}} = \frac{Y - 80}{\sqrt{64}} = \frac{Y - 80}{8} \xrightarrow{D} N(0,1) \quad \text{by CLT.}$$

Hence,

$$P(0.25 < Y/400) = P(100 < Y) = P(2.5 < (Y - 80)/8) = \texttt{1-pnorm(2.5)} = 0.0062.$$

If consider a continuity correction, then the probability is

$$P(Y > 100) = P(Y > 100.5) = P((Y - 80)/8 > 2.5625) = \texttt{1-pnorm(2.5625)} = 0.0052.$$

5.3.7. If Y is $b(100, \frac{1}{2})$, approximate the value of $P(Y = 50)$.

Solution.

$$\frac{Y - np}{\sqrt{np(1-p)}} = \frac{Y - 50}{\sqrt{25}} \xrightarrow{D} N(0,1) \quad \text{by CLT.}$$

Use a continuity correction to obtain

$$P(Y = 50) = P(49.5 \le Y \le 50.5) = P(-0.1 \le (Y - 50)/5 \le 0.1) = 0.080.$$

5.3.8. Let Y be $b(n, 0.55)$. Find the smallest value of n such that (approximately) $P(Y/n > \frac{1}{2}) \ge 0.95$.

Solution.

$$\frac{\sqrt{n}(Y/n - 0.55)}{\sqrt{(0.55)(0.45)}} \xrightarrow{D} N(0,1) \quad \text{by CLT.}$$

Hence,

$$0.95 \le P(Y/n > 1/2) = P\left(\frac{\sqrt{n}(Y/n - 0.55)}{\sqrt{(0.55)(0.45)}} > \frac{-0.05\sqrt{n}}{\sqrt{(0.55)(0.45)}}\right) = \Phi\left(\frac{0.05\sqrt{n}}{\sqrt{(0.55)(0.45)}}\right)$$

$$\Rightarrow \frac{0.05\sqrt{n}}{\sqrt{(0.55)(0.45)}} > 1.645 \Rightarrow n > \frac{1.645^2(0.55)(0.45)}{0.05^2} = 267.90,$$

which indicates that the smallest n is 268.

5.3.9. Let $f(x) = 1/x^2$, $1 < x < \infty$, zero elsewhere, be the pdf of a random variable X. Consider a random sample of size 72 from the distribution having this pdf. Compute approximately the probability that more than 50 of the observations of the random sample are less than 3.

Solution.

Let W is a Bernoulli trial:

$$W = \begin{cases} 1 & p = P(X < 3) \\ 0 & 1 - p, \end{cases}$$

where

$$p = \int_1^3 \frac{1}{x^2} dx = 1 - \frac{1}{3} = \frac{2}{3}.$$

Further let $Y = \sum_{i=1}^{72} W_i \sim b(72, p = 2/3)$, by CLT,

$$\frac{Y - np}{\sqrt{np(1-p)}} = \frac{Y - 48}{\sqrt{16}} = \frac{Y - 48}{4} \xrightarrow{D} N(0,1).$$

Hence, the desired probability is

$$P(Y \ge 50) = P((Y - 48)/4 \ge 0.5) = \texttt{pnorm(0.5, lower = F)} = 0.309.$$

If a continuity correction is used,

$$P(Y \ge 50) = P(Y \ge 50.5) = \texttt{pnorm(0.625, lower = F)} = 0.266.$$

5.3.10. Forty-eight measurements are recorded to several decimal places. Each of these 48 numbers is rounded off to the nearest integer. The sum of the original 48 numbers is approximated by the sum of these integers. If we assume that the errors made by rounding off are iid and have a uniform distribution over the interval $(-\frac{1}{2}, \frac{1}{2})$, compute approximately the probability that the sum of the integers is within two units of the true sum.

Solution.

Let $U_i \sim U(-0.5, 0.5)$. Then $E(U_i) = 0$, $\text{Var}(U_i) = [0.5 - (0.5)]/12 = 1/12$, which gives us

$$\sum_{i=1}^{48} U_i/\sqrt{48/12} = \sum_{i=1}^{48} U_i/2 \xrightarrow{D} N(0,1) \quad \text{by CLT.}$$

Then, the desired probability is

$$P\left(\left|\sum_{i=1}^{48} U_i\right| \le 2\right) = P\left(-2 \le \sum_{i=1}^{48} U_i \le 2\right) = P\left(-1 \le \sum_{i=1}^{48} U_i/2 \le 1\right) = 0.683.$$

5.3.11. We know that $\overline{X}$ is approximately $N(\mu, \sigma^2/n)$ for large n. Find the approximate distribution of $u(\overline{X}) = \overline{X}^3$, provided that $\mu \ne 0$.

Solution.

$$\overline{X} \text{ approx. } N(\mu, \sigma^2/n) \Leftrightarrow \sqrt{n}(\overline{X} - \mu) \xrightarrow{D} N(0, \sigma^2) \quad \text{by CLT.}$$

Let $g(x) = x^3$, which is continuous and differentiable at x ($g'(x) = 3x^2$). Then, by Delta method,

$$\sqrt{n}(\overline{X}^3 - \mu^3) \sim N(0, 9\mu^4\sigma^2) \quad \Leftrightarrow \quad \overline{X}^3 \text{ approx. } N(\mu^3, 9\mu^4\sigma^2/n).$$

5.3.12. Let $X_1, X_2, ..., X_n$ be a random sample from a Poisson distribution with mean μ. Thus, $Y = \sum_{i=1}^n X_i$ has a Poisson distribution with mean $n\mu$. Moreover, $X = Y/n$ is approximately $N(\mu, \mu/n)$ for large n. Show that $u(Y/n) = \sqrt{Y/n}$ is a function of Y/n whose variance is essentially free of μ.

Solution.

$$\overline{X} \text{ approx. } N(\mu, \mu/n) \quad \Leftrightarrow \quad \sqrt{n}(\overline{X} - \mu) \xrightarrow{D} N(0, \mu) \quad \text{by CLT.}$$

Let $g(x) = \sqrt{x}$, which is continuous and differentiable at x ($g'(x) = 1/(2\sqrt{x})$). Then, by Delta method,

$$\sqrt{n}(\sqrt{\overline{X}} - \sqrt{\mu}) \xrightarrow{D} N\left(0, \frac{1}{4}\right) \quad \Leftrightarrow \quad \sqrt{\overline{X}} \text{ approx. } N\left(\sqrt{\mu}, \frac{1}{4n}\right),$$

whose variance is free of μ.

5.3.13. Using the notation of Example 5.3.5, show that equation (5.3.4) is true.

Solution.

$$\frac{\widehat{p} - p}{\sqrt{\widehat{p}(1-\widehat{p})/n}} = \frac{\widehat{p} - p}{\sqrt{p(1-p)/n}} \frac{\sqrt{p(1-p)}}{\sqrt{\widehat{p}(1-\widehat{p})}} \xrightarrow{D} N(0,1) \quad \text{by Slutsky}$$

because

$$\frac{\widehat{p} - p}{\sqrt{p(1-p)/n}} \xrightarrow{D} N(0,1) \quad \text{by CLT,}$$

$$\frac{\sqrt{p(1-p)}}{\sqrt{\widehat{p}(1-\widehat{p})}} \xrightarrow{P} 1 \quad \text{by } g() \text{ and WLLN } (\widehat{p} \xrightarrow{P} E(X) = p).$$

5.3.14. Assume that $X_1, X_2, ..., X_n$ is a random sample from a $\Gamma(1, \beta)$ distribution. Determine the asymptotic distribution of $\sqrt{n}(X - \beta)$. Then find a transformation $g(\overline{X})$ whose asymptotic variance is free of β.

Solution.

Since $E(X) = \beta, \mathrm{Var}(X) = \beta^2$,

$$\sqrt{n}(\overline{X} - \beta) \xrightarrow{D} N(0, \beta^2) \quad \text{by CLT.}$$

Let $g(x) = \log x$, which is continuous at $x > 0$ and $g'(\beta) = 1/\beta$. Thus, by Delta Method,

$$\sqrt{n}(\log \overline{X} - \log \beta) \xrightarrow{D} N(0, \{g'(\beta)\}^2 \beta^2) = N(0, 1).$$

Hence, $\log \overline{X}$ is a transformation to make the asymptotic variance free of β.