This document contain's solution's to Exercises 5.29 to 5.36 from chapter 5, of statistical inference by Casella and Berger, focusing on convergence of random variables, WLLN, SLLN and CLT.

Some problem's are still in progress, and the document will be updated with their solution's soon.

5.29

$x_1, x_2, \ldots, x_{100} \overset{iid}{\sim} X$

$\mathbb{E}[X] = 1$

$sd(X) = 0.05 = \sigma$

$P(x_1 + x_2 + \ldots + x_{100} > 100.4)$?

Soln

we can use central limit theorem

were we can approximate

$$\mathbb{P}\left[\sqrt{n}\left(\frac{\bar{X}-\mu}{\sigma}\right) < c\right] < \mathbb{P}\left[Z < c\right]$$

$\forall c$, and $Z \sim N(0,1)$

Therefore

$\mathbb{P}(x_1 + x_2 + \ldots + x_{100} > 100.4)$

$$= \mathbb{P}\left(\overline{X}_n > \frac{100.4}{100}\right)$$

$$= \mathbb{P}\left(\overline{X}_n - 1 > \frac{100.4}{100} - 1\right)$$

$$= \mathbb{P}\left(\frac{\overline{X}_n - 1}{0.05} > \frac{0.4}{100.0.05}\right)$$

$$= \mathbb{P}\left(\frac{\overline{X}_n - 1}{0.05} > \frac{0.4}{5}\right)$$

$$= \mathbb{P}\left(\sqrt{100}\left(\frac{\overline{X}_n - 1}{0.05}\right) > \frac{4}{5}\right)$$

$$= 1 - \mathbb{P}\left(\sqrt{100}\left(\frac{\overline{X}_n - 1}{0.05}\right) \leq \frac{4}{5}\right)$$

for large n

$$\approx 1 - \mathbb{P}\left(Z \leq \frac{4}{5}\right)$$

$$\approx 1 - 0.7881 = 0.2219$$

$\Rightarrow$ $\mathbb{P}$(100 booklets weigh more than 100.4 ounces)

$$\boxed{\approx 0.2219}$$

5.30 if $\overline{x_1}$ & $\overline{x_2}$ are the mean's of two independend samples of size n from a popotation with variance σ^2. find a value of n so that

$$\mathbb{P}\left(|\overline{x}_1 - \overline{x}_2| < \frac{\sigma}{5}\right) \approx 0.99.$$

Justify your calculation's.

Soln

Assume $x_1, x_2, \ldots x_n$ iid X

$Var(X) = \sigma^2 \qquad E[X] = \mu$

$$\overline{X} = \frac{x_1 + x_2 + \cdots + x_n}{n} \Rightarrow Var(\overline{x}) = \frac{\sigma^2}{n}$$

$$E[\overline{x}] = \mu$$

$\overline{x_1}$ & $\overline{x_2}$ are two samples from $\overline{x}$

using CLT for large n

$$\overline{x_1} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

$$\overline{x_2} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

$$\Rightarrow \quad \bar{X}_1 - \bar{X}_2 \sim N\left(0, 2\frac{\sigma^2}{n}\right)$$

$$\Rightarrow \quad |\bar{X}_1 - \bar{X}_2| < \frac{\sigma}{5}$$

$$\Rightarrow \quad -\frac{\sigma}{5} < \bar{X}_1 - \bar{X}_2 < \frac{\sigma}{5}$$

$$\Rightarrow \quad \mathbb{P}\left(|\bar{X}_1 - \bar{X}_2| < \frac{\sigma}{5} \right)$$

$$\Rightarrow \quad \mathbb{P}\left(-\frac{\sigma}{5} < \bar{X}_1 - \bar{X}_2 < \frac{\sigma}{5} \right)$$

$$\Rightarrow \quad \mathbb{P}\left(\frac{-\sigma}{5 \cdot \sigma\sqrt{\frac{2}{n}}} < \frac{\bar{X}_1 - \bar{X}_2}{\sigma\sqrt{\frac{2}{n}}} < \frac{\sigma}{5 \cdot \sigma\sqrt{\frac{2}{n}}} \right)$$

$$0.99 \approx \mathbb{P}\left(-\frac{1}{5}\sqrt{\frac{n}{2}} < Z < +\frac{1}{5}\sqrt{\frac{n}{2}} \right)$$

$$0.99 \approx \mathbb{P}\left(Z < \frac{1}{5}\sqrt{\frac{n}{2}} \right) - \mathbb{P}\left(Z < -\frac{1}{5}\sqrt{\frac{n}{2}} \right)$$

$$0.99 \approx 1 - 2\,\mathbb{P}\left(Z < -\frac{1}{5}\sqrt{\frac{n}{2}} \right)$$

$$\mathbb{P}\left(Z < -\frac{1}{5}\sqrt{\frac{n}{2}}\right) \approx 0.005$$

$$-\frac{1}{5}\sqrt{\frac{n}{2}} = -2.576$$

$$\Rightarrow \sqrt{\frac{n}{2}} = 2.576 \times 5$$

$$\Rightarrow n = (2.576 \times 5)^2 \times 2$$

$$\Rightarrow n \approx 331.78 = 332$$

$$\boxed{n = 332}$$

5.31. $x_1, x_2, \ldots x_{100} \overset{iid}{\sim} X$

$E[X] = \mu \quad Var(X) = 9$

Find the limits b/w which $\bar{X} - \mu$ will lie with Proba at least 0.90. Use Both Chebychev's inequality and the CLT, and comment on each.

Sol^n

$$E[\bar{X}-\mu] = 0$$

$$\text{var}(\bar{X}-\mu] = \frac{\sigma^2}{100} = \frac{9}{100}$$

<u>using chebychev's inequality</u>

$$\mathbb{P}\left((\bar{X}-\mu)^2 \geq \alpha^2\right) \leq \frac{\mathbb{E}\left((\bar{X}-\mu)^2\right)}{\alpha^2}$$

$$\Rightarrow \mathbb{P}\left(|\bar{X}-\mu| \geq \alpha\right) \leq \frac{\text{var}((\bar{X}-\mu))}{\alpha^2}$$

$$\Rightarrow \mathbb{P}\left(|\bar{X}-\mu| \geq \alpha\right) \leq \frac{9}{100\alpha^2}$$

$$\Rightarrow 1 - \mathbb{P}\left(|\bar{X}-\mu| \geq \alpha\right) \geq 1 - \frac{9}{100\alpha^2}$$

$$= \mathbb{P}\left(|\bar{X}-\mu) < \alpha\right) \geq 1 - \frac{9}{100\alpha^2}$$

$$\Rightarrow \mathbb{P}\left(-\alpha < \bar{X}-\mu < \alpha\right) \geq 1 - \frac{9}{100\alpha^2}$$

$$\alpha = c\sigma = c \cdot \frac{3}{10}$$

$$\Rightarrow \quad \mathbb{P}\left(-\frac{3c}{10} < \bar{x}-\mu < \frac{3c}{10}\right) \geq 1-\frac{1}{c^2}$$

$$\geq 0.9$$

$$\Rightarrow \quad 1-\frac{1}{c^2} = 0.9$$

$$\Rightarrow \quad \frac{1}{c^2} = 0.1$$

$$\Rightarrow \quad c^2 = 10 \Rightarrow c = 3.1622$$

$$\mathbb{P}(-0.9486 < \bar{x}-\mu < 0.9486) \geq 0.9$$

<u>Using CLT</u>

$$\bar{X}-\mu \sim N\left(0, \frac{9}{100}\right)$$

$$\Rightarrow \quad \frac{\bar{X}-\mu}{3/10} \sim N(0,1)$$

$$\Rightarrow \quad \frac{(\bar{X}-\mu)}{3/10} \sim N(0,1)$$

$$0.9 = \mathbb{P}\left(-c < \frac{\bar{x}-\mu}{3/10} < c\right)$$

$c = 1.645$

$\Rightarrow \mathbb{P}\left(-1.645 \cdot \frac{3}{10} < \bar{x} - \mu < 1.645 \cdot \frac{3}{10}\right)$

$\Rightarrow \mathbb{P}\left(-0.4935 < \bar{x} - \mu < 0.4935\right)$

Bounds of Chebychev's inequality is high compared to CLT

$\Rightarrow$ CLT is far better in finding bounds than Chebychev's if n is large

5.32

$X_1, X_2, \dots$ be a seq of random variables that converges in prob to a constant a. Assume that $\mathbb{P}(X_i > 0) = 1$ $\forall i$

(a) Verify that the seq^n defined by

$Y_i = \sqrt{X_i}$ and $X_i' = \frac{a}{X_i}$ converge in Probability.

Sol^n the seq $X_1, X_2, \ldots X_n, X_{n+1}, \ldots$ converges in Probability to constant a

$$\Rightarrow \forall \varepsilon \quad \lim_{n \to \infty} \mathbb{P}(|X_n - a| > \varepsilon) = 0$$

we want to prove $\forall \varepsilon > 0$

$$\lim_{n \to \infty} \mathbb{P}(|\sqrt{X_n} - \sqrt{a}| > \varepsilon) = 0$$

$$\mathbb{P}(|\sqrt{X_n} - \sqrt{a}| > \varepsilon)$$

$$= \mathbb{P}(|\sqrt{X_n} - \sqrt{a}| \cdot |\sqrt{X_n} + \sqrt{a}| > \varepsilon |\sqrt{X_n} + \sqrt{a}|)$$

$$= \underbrace{\mathbb{P}(|X_n - a| > \varepsilon(\sqrt{X_n} + \sqrt{a})}_{= 0 \text{ as } n \to \infty})$$

$$\Rightarrow \mathbb{P}\left(|\sqrt{X_n} - \sqrt{a}| > \varepsilon\right) = 0$$

as $n \to$

(b) Use the result in Part (a) to prove the fact used in Example 5.5.18, that $\frac{\sigma}{S_n}$ converges in Prob to 1

Sol^n

we know that

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x}_n)^2$$

$$\mathbb{P}\left(|S_n^2 - \sigma^2| \geq \varepsilon\right) \leq \frac{\operatorname{var}(S_n^2)}{\varepsilon^2}$$

$$\Rightarrow \quad S_n^2 \longrightarrow \sigma^2 \quad \text{in Probability}$$

$$\Rightarrow \quad S_n \longrightarrow \sigma \quad \text{in Probability}$$

$\forall \varepsilon > 0 \quad \lim_{n \to \infty} \mathbb{P}(|S_n - \sigma| < \varepsilon) = 1$

let $\varepsilon > 0$

$$\mathbb{P}\left(\left| \frac{\sigma}{S_n} - 1 \right| < \varepsilon \right)$$

$$\Rightarrow \mathbb{P}\left(1 - \varepsilon < \frac{\sigma}{S_n} < 1 + \varepsilon \right)$$

$$\Rightarrow \mathbb{P}\left(\frac{1-\varepsilon}{\sigma} < \frac{1}{S_n} < \frac{1+\varepsilon}{\sigma} \right)$$

$$\Rightarrow \mathbb{P}\left(\frac{\sigma}{1+\varepsilon} < S_n < \frac{\sigma}{1-\varepsilon} \right)$$

$$\Rightarrow \mathbb{P}\left(\frac{\sigma}{1+\varepsilon} - \sigma < S_n - \sigma < \frac{\sigma}{1-\varepsilon} - \sigma \right)$$

$$\Rightarrow \mathbb{P}\left(\frac{-\sigma\varepsilon}{1+\varepsilon} < S_n - \sigma < \frac{\sigma\varepsilon}{1+\varepsilon} \right)$$

$$\Rightarrow \mathbb{P}\left(|S_n - \sigma| < \frac{\sigma \varepsilon}{1+\varepsilon} \right)$$

as $n \to \infty$ $\qquad = 1$

$\Rightarrow$ $\forall \varepsilon > 0$ we have

$$\lim_{n\to\infty} \mathbb{P}\left(\left| \frac{\sigma}{S_n} - 1 \right| < \varepsilon \right) = 1$$

$$\boxed{\Rightarrow \quad \frac{\sigma}{S_n} \xrightarrow{i.P} 1}$$

5.33 Let X_n be a seq of r.v's that converges in distribution to a random variable X. Let Y_n be a seq of r.v's with the property that for any finite number c;

$$\lim_{n\to\infty} \mathbb{P}(Y_n > c) = 1$$

Show that for any finite number c,

$$\lim_{n\to\infty} \mathbb{P}(X_n + Y_n > c) = 1$$

Sol

Sol^n

we have

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$$

$$\Rightarrow \lim_{n\to\infty} \mathbb{P}(X_n \le x) = \mathbb{P}(X_n \le x)$$

$$\Rightarrow 1 - \lim_{n\to\infty} \mathbb{P}(X_n \le x) = 1 - \mathbb{P}(X_n \le x)$$

$$\Rightarrow \lim_{n\to\infty} \mathbb{P}(X_n > x) = \mathbb{P}(X > x)$$

& we have

$$\lim_{n\to\infty} \mathbb{P}(Y_n > c) = 1$$

5.34 Let $X_1, X_2, \ldots$ be a random sample from a Population with mean μ and variance σ^2.

Show that

$$\mathbb{E}\left[\sqrt{n}\,\frac{(\bar{X}_n-\mu)}{\sigma}\right]=0$$

& $$\operatorname{Var}\left(\sqrt{n}\left(\frac{\bar{X}_n-\mu}{\sigma}\right)\right)=1$$

Sol^n $\mathbb{E}[\bar{X}]=\mu$ $\quad$ $\operatorname{Var}(\bar{X})=\frac{\sigma^2}{n}$

$$\mathbb{E}\left[\sqrt{n}\,\frac{(\bar{X}-\mu)}{\sigma}\right]$$

$$=\frac{\sqrt{n}}{\sigma}\left[E[\bar{X}]-E[\mu]\right]$$

$$=\frac{\sqrt{n}}{\sigma}\left[\mu-\mu\right]=0$$

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

$$\Rightarrow \frac{n}{\sigma^2}\operatorname{var}(\bar{X}-\mu)$$

$$\Rightarrow \frac{n}{\sigma^2}\operatorname{var}(\bar{X}) = \frac{n}{\sigma^2}\cdot\frac{\sigma^2}{n} = 1$$

5.35 Stirling's Formula

$$n! \approx \sqrt{2\pi}\, n^{n+\frac{1}{2}} \cdot e^{-n}$$

(a) Argue that, if $X_i \sim \exp(1)$, $i=1,2,\ldots$ all independent. then for every x,

$$P\left(\frac{\bar{X}_n - 1}{1/\sqrt{n}} \le x\right) \longrightarrow P(Z \le x)$$

$Z \sim N(0,1)$

Sol^n

$$X \sim \exp(1)$$

$$f_X(x) = e^{-x} \quad x > 0$$

$$E[X] = 1$$

$$\text{var}(X) = 1$$

then $\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$

$$\mathbb{E}[\bar{X}] = 1 \qquad \text{var}(\bar{X}) = \frac{1}{n}$$

Using CLT we know

$$\mathbb{P}\left(\frac{\bar{X}_n - 1}{1/\sqrt{n}} < x \right) \approx \mathbb{P}(Z < x)$$

(b) Show that differentiating both sides of the approximation in part (a) suggests.

Sol^n

$$\mathbb{P}\left(\frac{\bar{X} - 1}{1/\sqrt{n}} < x \right)$$

$$\Rightarrow \mathbb{P}\left(\bar{X} < \frac{x}{\sqrt{n}} + 1 \right)$$

$$\Rightarrow \mathbb{P}\left(\sum X_i < x\sqrt{n} + n \right)$$

we know that

$$X_i \sim \text{gamma}(1,1)$$

$$X_1+X_2+\cdots+X_n \sim \text{gamma}(n,1)$$

$$\Rightarrow f_{X_1+X_2+\cdots+X_n}(x) = \frac{1}{\Gamma(n)} x^{n-1} e^{-x}$$

$$\frac{d}{dx} F_{X_1+X_2\cdots+X_n}(x\sqrt{n}+n)$$

$$= f_{X_1+\cdots+X_n}(x\sqrt{n}+n)\cdot\sqrt{n}$$

$$= \frac{1}{\Gamma(n)}(x\sqrt{n}+n)^{n-1}\cdot e^{-(x\sqrt{n}+n)}\cdot\sqrt{n}$$

$$\frac{d}{dx} F_Z(x) = f_Z(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$\Rightarrow \frac{1}{\Gamma(n)}(x\sqrt{n}+n)^{n-1} \cdot e^{-(x\sqrt{n}+n)} \cdot \sqrt{n} \approx \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

as $n \to \infty$

Put $x=0$

$$\Rightarrow \frac{1}{\Gamma(n)} n^{n-1} \cdot e^{-n} \cdot \sqrt{n} \approx \frac{1}{\sqrt{2\pi}}$$

$$\Rightarrow \frac{n^n}{n \cdot (n-1)!} \cdot e^{-n} \sqrt{n} \approx \frac{1}{\sqrt{2\pi}}$$

$$\Rightarrow n! \approx \sqrt{2\pi} \cdot n^{n+\frac{1}{2}} \cdot e^{-n}$$

$$\Rightarrow \boxed{n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n}$$

5.36 Given that $N=n$, the conditional distribution of Y is χ^2_{2n}.

The unconditional distribution of N is Poisson(θ)

(a) Calculate $E[Y]$ and $\mathrm{var}(Y)$

Sol^n

$$N \sim \text{Poisson}(\theta)$$

$$\mathbb{P}(N=n) = \frac{e^{-\theta} \cdot \theta^n}{n!} \qquad n=0,1,2,\ldots$$

Given

$$\mathbb{P}(Y \mid N=n) \text{ is } \chi^2_{2n}$$

$$\Rightarrow \quad Y \mid N=n \sim \chi^2_{2n}$$

$$Y \mid N=n \sim \text{gamma}(n, 2)$$

$$N \sim \text{Poisson}(\theta)$$

Using Theorem 4.4.3

$$\mathbb{E}[Y] = \mathbb{E}[E[Y|N]]$$

$$= \mathbb{E}[2N]$$

$$\boxed{E[Y] = 2\theta}$$

Using Theorem 4.4.7

$$\text{var}(Y) = \mathbb{E}[\text{var}(Y|N)] + \text{var}(E[Y|N])$$

$$= \mathbb{E}[4N] + \text{var}(2N)$$

$$= 4\theta + 2\theta = 6\theta$$

$$\boxed{\text{var}(Y) = 6\theta}$$

(b) Show that as $\theta \to \infty$

$$\frac{(Y - E[Y])}{\sqrt{var(Y)}} \longrightarrow n(0,1)$$

Sol^n

$$Y|N \sim \chi^2_{2N}$$

$$\Rightarrow Y|N \sim \underbrace{\chi^2_1 + \chi^2_1 + \cdots \chi^2_1}_{2N \text{ term's.}}$$

$$\mathbb{P}(Y=y) = \sum_{n=0}^{\infty} \mathbb{P}(Y=y, N=n)$$

$$= \sum_{n=0}^{\infty} \mathbb{P}(Y=y \mid N=n)\, P(N=n)$$

$$= \sum_{n=0}^{\infty} \frac{1}{\Gamma(n)\, 2^n}\, y^{n-1} \cdot e^{-\frac{y}{2}} \cdot \frac{e^{-\theta} \cdot \theta^n}{n!}$$

$$= \frac{1}{y} \cdot e^{-\theta} \cdot e^{-\frac{y}{2}} \sum_{n=0}^{\infty} \frac{1}{\Gamma(n)\, 2^n} \cdot y^n \cdot \frac{\theta^n}{n!}$$

$$= \frac{e^{-\left(\theta+\frac{y}{2}\right)}}{y} \sum_{n=0}^{\infty} \left(\frac{y\theta}{2}\right)^{n} \cdot \frac{1}{\Gamma(n)\, n!}$$

(this method is not working)

$$\frac{Y - E[Y]}{\operatorname{var}(Y)} = \frac{Y - 2\theta}{\sqrt{6\theta}}$$

5.44 Let X_i's, $i = 1, 2, \ldots$ be independent Bernoulli (P) random variable and let

$$Y_n = \frac{1}{n} \sum_{i=1}^{n} x_i$$

(a) Show that $\sqrt{n}\,(Y_n - P) \longrightarrow n(0, P(1-P))$ in distribution

Sol^n using CLT

$$E[Y_n] = E[X] = nP$$

$$Var[Y_n] = \frac{Var(X)}{n} = \frac{\cancel{n}P(1-P)}{\cancel{n}}$$

$$= P(1-P)$$

$\Rightarrow$ By CLT

$$\sqrt{n}\left(\frac{Y_n - nP}{\sqrt{P(1-P)}}\right) \longrightarrow N(0,1)$$

$$\Rightarrow \sqrt{n}\,(Y_n - nP) \longrightarrow \sqrt{P(1-P)}\; N(0,1)$$

$$\longrightarrow N(0, P(1-P))$$