Bryant Willoughby

University of Michigan, Dept of Statistics

Stats 510, Instructor: Long Nguyen

Homework 6

Dec 3, due by 11:59pm Dec 12, 2025

1. Do problems 5.3, 5.6, 5.8, 5.10.

2. Construct a sequence of non-identically distributed random variables $\{X_n\}_{n=1}^{\infty}$ and a random variable Y in the following settings. (Specify both X_n and Y as functions of a sample space S of your choice, and briefly justify your answer).

   (i) X_n converges to Y almost surely, where $Y \sim \text{Uniform}(0, 1)$.

   (ii) X_n converges to Y in probability, where $Y \sim$ Uniform(0,1), but X_n does not converge to Y almost surely.

   (iii) For each n, X_n is a binary random variable that takes two possible values (0 or 1), and so is Y, and that X_n converges to Y in distribution, but not in probability.

3. Let $X_1, \ldots, X_n, \ldots$ be a sequence of independent random variables with a shared mean μ. X_i has variance σ_i for all i. State a sufficient condition according to which you can show that the weak law of large numbers remains true, i.e., $\bar{X}_n := (1/n)(X_1 + \ldots + X_n)$ tends to μ in probability.

4. Let $X_1, X_2, \ldots$ be a sequence of random variables that converges in probability to a constant a. Assume that $P(X_i > 0) = 1$ for all i. Show that the sequences defined by $Y_i = \sqrt{X_i}$ and $Y_i' = a/X_i$ converge in probability.

5. Let $\{X_n\}$ be sequence of random variables that converges in distribution to a random variable X. Let $\{Y_n\}$ be a sequence of random variables such that for any finite c, $\lim_{n\to\infty} P(Y_n > c) = 1$. Show that for any finite c,

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

1. Do problems 5.3, 5.6, 5.8, 5.10.

5.3 Let $X_1, \ldots, X_n$ be iid random variables with continuous cdf F_X, and suppose $EX_i = \mu$. Define the random variables $Y_1, \ldots, Y_n$ by

$$Y_i = \begin{cases} 1 & \text{if } X_i > \mu \\ 0 & \text{if } X_i \le \mu. \end{cases}$$

Find the distribution of $\sum_{i=1}^n Y_i$.

Note that $Y_i \sim$ Bernoulli with

$p = P(Y_i = 1) = P(X_i > \mu) = P(X_i \ge \mu) = 1 - F(\mu) \ \forall i$

Then, $Y_1, \ldots, Y_n \overset{iid}{\sim}$ Bernoulli$(p = 1 - F(\mu))$.

Thus, $\sum_{i=1}^{n} Y_i \sim$ Binomial$(n, p = 1 - F(\mu))$

5.6 If X has pdf $f_X(x)$ and Y, independent of X, has pdf $f_Y(y)$, establish formulas, similar to (5.2.3), for the random variable Z in each of the following situations.

(a) $Z = X - Y$

(b) $Z = XY$

(c) $Z = X/Y$

a) Define transformation $\left| \begin{array}{l} z = x - y \\ w = x \end{array}\right.$ $\Rightarrow$ inverse transformation $\left| \begin{array}{l} x = w \\ y = w - z \end{array}\right.$

we have a one-to-one transformation : $(x, y) \to (w, w - z)$

Next compute $J = \left[\dfrac{\partial(x,y)}{\partial(z,w)}\right] = \begin{bmatrix} \frac{\partial x}{\partial z} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial z} & \frac{\partial y}{\partial w} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix} \Rightarrow |J| = 0 + 1 = |1| = 1.$

using independence,

$f_{Z,W}(z,w) = f_X(x) f_Y(y) \cdot |J| = f_X(w)\, f_Y(w - z) \cdot |J|$

$= f_X(w)\, f_Y(w - z)$

Now marginalize (out) w to find

$$\boxed{f_Z(z) = \int_{-\infty}^{\infty} f_X(w)\, f_Y(w - z)\, dw}$$

b) Define transformation $\left| \begin{array}{l} z = xy \\ w = x \end{array}\right.$ $\Rightarrow$ inverse transformation $\left| \begin{array}{l} x = w \\ y = z/w = w^{-1} z \end{array}\right.$

$J = \begin{bmatrix} \frac{\partial x}{\partial z} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial z} & \frac{\partial y}{\partial w} \end{bmatrix} = \begin{pmatrix} 0 & 1 \\ 1/w & -z/w^2 \end{pmatrix} = -1/w \Rightarrow |J| = 1/w$

Then,

$f_{Z,W}(z,w) = f_X(w)\, f_Y(z/w) \cdot \frac{1}{w}$

$$\Rightarrow \boxed{f_Z(z) = \int_{-\infty}^{\infty} f_X(w)\, f_Y(z/w)\, \left|\tfrac{1}{w}\right|\, dw}$$

c) For $\left| \begin{array}{l} z = x/y \\ w = x \end{array}\right.$ $\Rightarrow \left| \begin{array}{l} y = w/z \\ x = w \end{array}\right.$ and $|J| = \begin{vmatrix} 0 & 1 \\ -w/z^2 & 1/z \end{vmatrix} = w/z^2$

So,

$f_{Z,W}(z,w) = f_X(w) f_Y(w/z) \cdot |w/z^2|$

$$\Rightarrow \boxed{f_Z(z) = \int_{-\infty}^{\infty} |w/z^2|\, f_X(w)\, f_Y(w/z)\, dw}$$

5.8 Let $X_1, \ldots, X_n$ be a random sample, where $\bar{X}$ and S^2 are calculated in the usual way.

(a) Show that

$$S^2 = \frac{1}{2n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} (X_i - X_j)^2.$$

Assume now that the X_is have a finite fourth moment, and denote $\theta_1 = EX_i, \theta_j = E(X_i - \theta_1)^j, j = 2, 3, 4.$

(b) Show that Var $S^2 = \frac{1}{n}(\theta_4 - \frac{n-3}{n-1}\theta_2^2)$.

(c) Find Cov$(\bar{X}, S^2)$ in terms of $\theta_1, \ldots, \theta_4$. Under what conditions is Cov$(\bar{X}, S^2) = 0$?

a) $\dfrac{1}{2n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} (X_i - X_j)^2 = \dfrac{1}{2n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} (X_i - \bar{X} + \bar{X} - X_j)^2$

where $(X_i - \bar{X} + \bar{X} - X_j) = (X_i - \bar{X}) - (X_j - \bar{X})$

$= \dfrac{1}{2n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} [(X_i - \bar{X})^2 + 2(X_i - \bar{X})(X_j - \bar{X}) + (X_j - \bar{X})^2]$

$= \dfrac{1}{2n(n-1)} \left[\sum_{i=1}^{n} n(X_i - \bar{X})^2 - 2\sum_{i=1}^{n}(X_i - \bar{X}) \underbrace{\sum_{j=1}^{n}(X_j - \bar{X})}_{0} + n \sum_{j=1}^{n} (X_j - \bar{X})^2 \right]$

PS since $\sum_{j=1}^{n}(X_j - \bar{X}) = \sum_j X_j - n\bar{X} = \sum_j X_j - \sum_j X_j = 0$

$= \dfrac{n}{2n(n-1)} \sum_{i=1}^{n} (X_i - \bar{X})^2 + \dfrac{n}{2n(n-1)} \sum_{j=1}^{n} (X_j - \bar{X})$

$= \dfrac{1}{2(n-1)} \cdot 2 \sum_{i=1}^{n} (X_i - \bar{X})^2$

$= \dfrac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2 = S^2.$ ∎

b) Assume now that the X_i's have a finite fourth moment, and denote

$\theta_1 = EX_i,\ \theta_j = E(X_i - \theta_1)^j,\ j = 2, 3, 4$

we proceed via proof by induction

We can assume $\theta_1 = EX_i = 0$ w/o loss of generality b/c both the sample variance and central moments θ_2, θ_4 are translation invariant, i.e. $\begin{cases} \theta_2 = EX_i^2 \\ \theta_4 = EX_i^4 \end{cases}$

<u>Base case</u>: First show that $\overbrace{\mathrm{Var}(S^2) = \frac{1}{n}[\theta_4 - \frac{n-3}{n-1}\theta_2^2]}^{(*)}\Big|_{n=4} = \frac{1}{4}(\theta_4 - \frac{1}{3}\theta_2^2)$

From a, $S^2 = \dfrac{1}{2n(n-1)} \sum_i \sum_j (X_i - X_j)^2 \Big|_{n=4} = \dfrac{1}{24} \sum_{i=1}^{4} \sum_{j=1}^{4} (X_i - X_j)^2$

we want to compute $\mathrm{Var}(S^2) = E[(S^2)^2] - (E(S^2))^2$

Assuming $X_i \overset{iid}{\sim}$ with $EX_i = 0$, compute $E(S^2)$,

expanding $S^2 = \left| \begin{array}{l} i = j : (X_i - X_j)^2 = 0 \ \ (4 \text{ such terms}) \\ i \ne j : E[(X_i - X_j)^2] = E(X_i^2) + E(X_j^2) - 2EX_iX_j \\ \qquad = \theta^2 + \theta^2 - 0 \\ \qquad = 2\theta^2 \ \ (4^2 - 4 = 12 \text{ such terms}) \end{array}\right.$

$\Rightarrow E(S^2) = \frac{1}{24}(12)(2\theta^2) = \theta^2.$

Now compute $E[(S^2)^2]$,

expanding $(S^2)^2 = \dfrac{1}{576} \sum_{i,j,k,l} (X_i - X_j)^2 (X_k - X_l)^2 =$

i) $i = j$ or $k = l \to$ expression $= 0$
$(4 \times 16 + 4 \times 16 - 4^2 = 112$ terms$)$

ii) $i \ne j : (X_i - X_j)^4 = X_i^4 - 4X_i^3X_j + 6X_i^2X_j^2 - 4X_iX_j^3 + X_j^4$
(24 terms) where $E(X_i) = 0$ and ind. kills odd moments:
$E[(X_i - X_j)^4] = 2\theta_4 + 6\theta_2 = 2(\theta_4 + 3\theta_2^2)$

iii) some distinct indices of form: $(X_i - X_j)^2 (X_j - X_k)^2$;
(96 terms)
$(X_i - X_j)^2 = X_i^2 + X_j^2 - 2X_iX_j$
$(X_j - X_k)^2 = X_j^2 + X_k^2 - 2X_jX_k$
$\Rightarrow E[(X_i - X_j)^2 (X_j - X_k)^2] = \theta_4 + 3\theta_2$

iv) all distinct indices: $(X_i - X_j)^2 (X_k - X_l)^2$
(24 terms) $\Rightarrow E[(X_i - X_j)^2 (X_k - X_l)^2] = E[(X_i - X_j)^2]\, E[(X_l - X_k)^2]$ via ind.
$= 2\theta_2^2\, 2\theta_2^2$
$= 4\theta_2^2$

So, $E[(S^2)^2] = \frac{1}{576}[24 \cdot (ii) + 96 \cdot (iii) + 24 (iv)] = \frac{1}{4}\theta_4 + \frac{11}{12}\theta_2^2$

and $V(S^2) = E[(S^2)^2] - (E(S^2))^2$

$= \left(\frac{1}{4}\theta_4 + \frac{11}{12}\theta_2^2\right) - \theta_2^2$

$= \frac{1}{4}(\theta_4 - \frac{1}{3}\theta_2^2)$

<u>Induction Step</u>. Assume that (*) holds for n and establish it for $n+1$.

Let $S_n^2 :=$ variance based on n observations.

Then, $S_{n+1}^2 = \dfrac{1}{2n(n+1)} \left[\underbrace{\sum_{i=1}^{n} \sum_{j=1}^{n} (X_i - X_j)^2}_{A} + 2 \underbrace{\sum_{k=1}^{n} (X_k - X_{n+1})^2}_{B} \right]$

where B: total squared difference btwn the new obs. X_{n+1} and each of the n previous ones

$= \dfrac{1}{2n(n+1)}[A + 2B]$

Then $\mathrm{var}(S_{n+1}^2) = \left(\frac{1}{2n(n+1)}\right)^2 \mathrm{var}(A + 2B)$

where $\mathrm{var}(A + 2B) = \mathrm{var}(A) + 4\,\mathrm{var}(B) + 4\,\mathrm{cov}(A, B)$.

By the induction hypothesis,

$S_n^2 = \dfrac{1}{2n(n-1)} A \Rightarrow A = 2n(n-1) S_n^2$

$\Rightarrow \mathrm{var}(A) = [2n(n-1)]^2 [\mathrm{var}(S_n^2) = \frac{1}{n}(\theta_4 - \frac{n-3}{n-1}\theta_2^2)]$

$= 4n(n-1)^2 [\theta_4 - \frac{n-3}{n-1}\theta_2^2]$

Given $B = \sum_{k=1}^{n} (X_k - X_{n+1})$, assuming $X_k \overset{iid}{\sim}$ $k = 1, \ldots, n$ and X_{n+1} independent of other X_k,

$\mathrm{var}(B) = \sum_{k=1}^{n} \mathrm{var}((X_k - X_{n+1})^2) + \sum_{k \ne l} \mathrm{cov}[(X_k - X_{n+1})^2, (X_l - X_{n+1})^2]$

$= n(n+1)\theta_4 - n(n-3)\theta_2^2$

Similarly, $\mathrm{cov}(A, B) = E(AB) - E(A)E(B)$

via tedious calculation reduces to

$= 2n(n-1)[\theta_4 - \theta_2^2]$

Then, after simplification,

$\mathrm{var}(S_{n+1}^2) = \dfrac{1}{n+1}\left[\theta_4 - \frac{n-2}{n}\theta_2^2\right],$

establishing the induction and verifying the result. ∎

c) $\mathrm{cov}(\bar{X}, S^2) = \mathrm{cov}\left(\frac{1}{n}\sum_k X_k, \frac{1}{2n(n-1)} \sum_{i,j} (X_i - X_j)^2\right)$

$= E(\bar{X} \cdot S^2) - E\bar{X} \cdot ES^2$

where $E\bar{X} = 0$ by assuming $\theta_1 = EX_i = 0$

$= E(\bar{X} \cdot S^2)$

$= \dfrac{1}{2n^2(n-1)} E\left[\left(\sum_{k=1}^{n} X_k\right)\left(\sum_{i=1}^{n}\sum_{j=1}^{n}(X_i - X_j)^2\right)\right]$

If $k \ne i$ and $k \ne j$, then X_k is independent of both X_i and X_j and since $E(X_k) = 0$, we get

$E(X_k (X_i - X_j)^2) = E(X_k) \cdot E[(X_i - X_j)^2] = 0$

So only terms where $k = i$ or $k = j$ contribute non-zero expectation

- $\forall$ fixed pair (i,j), there are 2 k values which contribute ($k = i$ or $k = j$)
- there are $n(n-1)$ total values of (i,j) with $i \ne j$:

<u>P</u>: For $i, j \in \{1, \ldots, n\}$,
total # pairs (i,j) w/o restriction $= n \times n = n^2$
There are n-such diagonal pairs where $i = j$
e.g. $(1,1), (2,2)$, etc.
$\Rightarrow$ # pairs where $i \ne j = n^2 - n = n(1-n)$. ∎

$\Rightarrow$ # nonzero expectation terms: $2n(n-1)$.

Thus, $\mathrm{cov}(\bar{X}, S^2) = \dfrac{2n(n-1)}{2n^2(n-1)} EX_i(X_i - X_j)^2$; computing a representative term:

$E[X_i(X_i - X_j)^2] = E[X_i(X_i^2 - 2X_iX_j + X_j^2)]$

where $\begin{cases} E(X_i^3) = \theta_3 \\ E(X_i^2 X_j) = E(X_i^2)E(X_j) = \theta_2 \cdot 0 = 0 \\ E(X_i X_j^2) = 0 \end{cases}$

$= \theta^3$; the same goes for $E[X_j(X_i - X_j)^2]$ since $X_i \sim X_j$

So, $\mathrm{cov}(\bar{X}, S^2) = \dfrac{2n(n-1)}{2n^2(n-1)} \cdot \theta^3 = \dfrac{\theta^3}{n}.$

$\Rightarrow \mathrm{cov}(\bar{X}, S^2) = 0$ i.f.f. $\theta_3 = 0$, i.e. the distribution of X_i is symmetric abt the mean.

(recall: $\theta_3 = E(X_i^3)$, i.e. the third central moment and the numerator of skewness)

5.10 Let $X_1, \ldots, X_n$ be a random sample from a $\mathrm{n}(\mu, \sigma^2)$ population.

(a) Find expressions for $\theta_1, \ldots, \theta_4$, as defined in Exercise 5.8, in terms of μ and σ^2.

(b) Use the results of Exercise 5.8, together with the results of part (a), to calculate $\operatorname{Var} S^2$.

(c) Calculate $\operatorname{Var} S^2$ a completely different (and easier) way: Use the fact that $(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$.

Let $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

a) $\theta_1 = E(X_i) = \mu$

$\theta_2 = \operatorname{Var}(X_i) = \sigma^2$

$\theta_3 = E(X_i - \mu)^3 = E(X_i-\mu)^2(X_i - \mu)$

FS recall (Stein's lemma): $Eg(X)(X-\theta) = \sigma^2 E g'(X)$

$= 2\sigma^2 \underbrace{E(X_i - \mu)}_{E(X_i) - E(\mu) = 0}$

$= 0$, a direct result of the symmetry of the normal dist.

$\theta_4 = E(X_i - \mu)^4 = E(X_i - \mu)^3(X_i - \mu) = 3\sigma^2 E(X_i - \mu)^2 = 3\sigma^4$

b) $\operatorname{Var} S^2 = \frac{1}{n}\left(\theta_4 - \frac{n-3}{n-1}\theta_2^2\right)$

$= \frac{1}{n}\left(3\sigma^4 - \frac{n-3}{n-1}\sigma^4\right)$

$= \frac{\sigma^4}{n}\left(3 - \frac{n-3}{n-1}\right)$

$= \frac{\sigma^4}{n}\left(\frac{3(n-1) - (n-3)}{n-1}\right)$

$= \frac{\sigma^4}{n}\left(\frac{2n}{n-1}\right)$

$= \frac{2\sigma^4}{n-1}$

c) Use the fact that $(n-1)\frac{S^2}{\sigma^2} \sim \chi^2_{n-1}$

and $\underbrace{\operatorname{Var}(X \sim \chi^2_k) = 2k}_{\text{FS}} \Rightarrow \operatorname{Var}\chi^2_{n-1} = 2(n-1)$

so, $\operatorname{Var}\left(\frac{(n-1)S^2}{\sigma^2}\right) = 2(n-1) \Rightarrow \left(\frac{(n-1)^2}{\sigma^4}\right)\operatorname{Var}S^2 = 2(n-1)$

and hence

$\operatorname{Var}S^2 = \frac{2(n-1)}{(n-1)^2/\sigma^4} = \frac{2\sigma^4}{n-1}$. ▪

2. Construct a sequence of non-identically distributed random variables $\{X_n\}_{n=1}^{\infty}$ and a random variable Y in the following settings. (Specify both X_n and Y as functions of a sample space S of your choice, and briefly justify your answer).

(i) X_n converges to Y almost surely, where $Y \sim \text{Uniform}(0, 1)$.

(ii) X_n converges to Y in probability, where $Y \sim \text{Uniform}(0,1)$, but X_n does not converge to Y almost surely.

(iii) For each n, X_n is a binary random variable that takes two possible values (0 or 1), and so is Y, and that X_n converges to Y in distribution, but not in probability.

Fix a probability space $S = [0,1]$; let $\omega \in [0,1]$ be the sample point and define all RVs as functions $X_n(\omega)$ and $Y(\omega)$

i) Let $Y(\omega) = \omega \sim \text{Uniform}(0,1)$

$X_n(\omega) = Y(\omega) + \frac{1}{n} = \omega + \frac{1}{n} \sim \text{Uniform}(\frac{1}{n}, 1 + \frac{1}{n})$;

notice that X_n is not identically distributed

$\forall \omega \in [0,1]$,

$X_n(\omega) = \omega + \frac{1}{n} \to \omega = Y(\omega)$ as $n \to \infty$

$\Rightarrow \underline{X_n \xrightarrow{a.s.} Y}$. ▪

ii) Let $Y \sim \text{Uniform}(0,1)$ and define $X_n := \begin{cases} Y, & \text{with prob. } 1 - \frac{1}{n} \\ Y+1, & \text{with prob. } \frac{1}{n} \end{cases}$

Each X_n is non-identically distributed since the mixture prob.'s differ

To show convergence in probability, we want to compute

$P(|X_n - Y| > \varepsilon)$.

In this construction $\begin{cases} \text{w/ prob. } 1-\frac{1}{n},\ X_n = Y \Rightarrow |X_n - Y| = 0 \\ \text{w/ prob. } \frac{1}{n},\ X_n = Y+1 \Rightarrow |X_n - Y| = 1 \end{cases}$

So the event $|X_n - Y| \geq \varepsilon$ only occurs when $X_n = Y+1$ w/ prob. $1/n$.

Hence: $P(|X_n - Y| \geq \varepsilon) = \begin{cases} \frac{1}{n} & \text{if } \varepsilon \leq 1 \\ 0 & \text{if } \varepsilon > 1 \end{cases}$

Thus, $\forall \varepsilon > 0$, $P(|X_n - Y| \geq \varepsilon) \leq \frac{1}{n} \to 0$ as $n \to \infty$

$\Rightarrow X_n \xrightarrow{P} Y$. ▫

Define the event where X_n differs from Y, i.e.

$A_n = \{\omega : X_n(\omega) \neq Y(\omega)\}$ s.t. $P(A_n) = \frac{1}{n}$

we examine whether $X_n(\omega) \neq Y(\omega)$ happens infinitely often (i.o.):

★ $\sum_{n=1}^{\infty} P(A_n) = \sum_{n=1}^{\infty} \frac{1}{n} = \infty$, i.e. the harmonic series diverges

Since events A_n are independent,

citing (second) Borel-Cantelli lemma: | If $\sum P(A_n) = \infty$ and A_n - independent then $P(A_n \text{ i.o.}) = 1$

Thus, $P(X_n \neq Y \text{ i.o.}) = 1 \Rightarrow P(\lim_{n\to\infty} |X_n - Y| = 0)$

$= P(\lim_{n \to \infty} X_n = Y) = 0$

$\Rightarrow X_n \not\xrightarrow{a.s.} Y$. ▪

iii) Define $X_n \sim \text{Bernoulli}(p_n)$ where $p_n = \frac{1}{2} + \frac{(-1)^n}{2n}$. Let $Y \sim \text{Bernoulli}(1/2)$.

Show: $X_n \xrightarrow{d} Y \sim \text{Bernoulli}(1/2)$:

$\forall x < 0$, $F_{X_n}(x) = 0 = F_Y(x)$

For $x \in [0,1)$:

$F_{X_n}(x) = P(X_n = 0) = 1 - p_n = \frac{1}{2} - \frac{(-1)^n}{2n} \to \frac{1}{2} = F_Y(x)$.

For $x \geq 1$: $F_{X_n}(x) = 1 = F_Y(x)$. ▫

Assuming X_n is independent of Y, then

$P(X_n \neq Y) = P(X_n = 0, Y = 1) + P(X_n = 1, Y = 0)$

$= P(X_n = 0)P(Y=1) + P(X_n = 1)P(Y = 0)$

$= (1 - p_n)\frac{1}{2} + p_n \cdot \frac{1}{2}$

$= \left(1 - \left(\frac{1}{2} + \frac{(-1)^n}{2n}\right)\right)\cdot\frac{1}{2} + \left(\frac{1}{2} + \frac{(-1)^n}{2n}\right)\cdot \frac{1}{2}$

$= \frac{1}{4} - \frac{(-1)^n}{4n} + \frac{1}{4} + \frac{(-1)^n}{4n} = \frac{1}{2}$

So,

$P(|X_n - Y| \geq \varepsilon = 1/2) = P(X_n \neq Y) = \frac{1}{2} \not\to 0$ as $n \to \infty$

$\Rightarrow X_n \not\xrightarrow{P} Y$. ▪

3. Let $X_1, \ldots, X_n, \ldots$ be a sequence of independent random variables with a shared mean μ. X_i has variance σ_i for all i. State a sufficient condition according to which you can show that the weak law of large numbers remains true, i.e., $\bar{X}_n := (1/n)(X_1 + \ldots + X_n)$ tends to μ in probability.

<u>WLLN</u>: = Let $X_1, \ldots, X_n$ iid RV's with $\begin{array}{|l} EX_i = \mu \\ \operatorname{Var}X_i = \sigma^2 \end{array}$

then $\bar{X}_n = \frac{1}{n}(X_1 + \cdots + X_n) \xrightarrow{P} \mu$.

A sufficient condition to show WLLN holds true is a finite variance: $\sigma_i^2 < \infty$.

Applying Chebyshev's Inequality:

$\forall \varepsilon > 0$, $P(|\bar{X}_n - \mu| \geq \varepsilon) \leq \frac{\operatorname{Var}(\bar{X}_n)}{\varepsilon^2} = \frac{\sigma^2/n}{\varepsilon^2} \to 0$ as $n \to \infty$

$\Rightarrow \bar{X}_n \xrightarrow{P} \mu$. ▪

4. Let $X_1, X_2, \ldots$ be a sequence of random variables that converges in probability to a constant a. Assume that $P(X_i > 0) = 1$ for all i. Show that the sequences defined by $Y_i = \sqrt{X_i}$ and $Y_i' = a/X_i$ converge in probability.

Let $X_1, X_2, \ldots \xrightarrow{P} a \Leftrightarrow X_i \xrightarrow{P} a$.

Assume that $P(X_i > 0) = 1$ for all i.

In subsequent derivations, I will use the

<u>continuous mapping thm</u> (CMT): = If $X_i \xrightarrow{P} a$ and g is continuous at a then $g(X_i) \xrightarrow{P} g(a)$

and the <u>lemma</u>: = If $\begin{array}{|l} X_n \xrightarrow{P} X \\ Y_n \xrightarrow{P} b: \text{constant} \end{array}$ then $X_n + Y_n \xrightarrow{P} aX + b$

Given $X_i \xrightarrow{P} a$,

- $g(x) = \sqrt{x}$ is a continuous function on $(0, \infty)$
- Since $P(X_i > 0) = 1$, $\sqrt{x}$ is well-defined and continuous at a

Thus, by the CMT,

$Y_i = \sqrt{X_i} \xrightarrow{P} \sqrt{a}$

Next, since $X_i \xrightarrow{P} a > 0$,

$Y_i' = \frac{a}{X_i} \xrightarrow{P} \frac{a}{g(x_i)} = \frac{a}{a} = 1$. ▪

5. Let $\{X_n\}$ be sequence of random variables that converges in distribution to a random variable X. Let $\{Y_n\}$ be a sequence of random variables such that for any finite c, $\lim_{n\to\infty} P(Y_n > c) = 1$. Show that for any finite c,

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

Let $X_n \overset{d}{\Rightarrow} X$. Let Y_n : sequence of RV's s.t. $\forall$ finite c, $\lim_{n\to\infty} P(Y_n > c) = 1$.

Show that for any finite c, $\lim_{n\to\infty} P(X_n + Y_n > c) = 1$:

Fact: = If one sequence Y_n diverges to $+\infty$ in prob., then adding any tight or convergent sequence cannot stop the sum from diverging.

Formally,

$Y_n \xrightarrow{P} +\infty$ means

$\forall M \in \mathbb{R}$, $P(Y_n > M) \to 1$.

$X_n \xrightarrow{d} X$ implies tightness of $\{X_n\}$, i.e.

$\forall \varepsilon > 0 \ \exists M$ s.t.

$\sup_n P(|X_n| > M) < \varepsilon$, recall: an upper bound b of S is a $\underbrace{\text{supremum (of } S)}_{=:\ \sup b}$ if for $z \in$ partially ordered set $(P, \leq)$,

$z \geq x$ for all $x \in S$.

In general, all sequences that converge in distribution are tight.

a lower bound a of S is called an infimum (of S):

Inf a := for $y \in (P, \leq)$,

$y \leq x$ for all $x \in S$.

Proof. want to show

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

consider the event decomposition, letting $M \in \mathbb{R}$:

$\{X_n + Y_n > c\} \supseteq \{Y_n > c + M\} \cap \{|X_n| \leq M\}$

citing $P(A \cap B) = P(A) - P(A \cap B^c) \geq P(A) - P(B^c)$

$\geq P(Y_n > c + M) - P(|X_n| > M)$

controlling $P(|X_n| > M)$ via tightness:

since $X_n \xrightarrow{d} X$, $\forall \varepsilon > 0$, $\exists M$ s.t.

$\sup_n P(|X_n| > M) < \varepsilon$.

Fix this M.

Thus, the 'bad' tail event of X_n has prob. at most ε.

By assumption,

$P(Y_n > c + M) \to 1$ as $n \to \infty$.

For sufficiently large n,

$P(X_n + Y_n > c) \geq P(Y_n > c + M) - \varepsilon$

$\Rightarrow \liminf_{n\to\infty} P(X_n + Y_n > c) \geq 1 - \varepsilon$

where $\varepsilon > 0$ was arbitrary

$= 1$

Thus, $\lim_{n\to\infty} P(X_n + Y_n > c) = 1$. ■