End of W5 lectures.

But then $\mathbb{E}(X) \geq \mathbb{E}(\delta Z) = \delta \mathbb{E}(Z)$

since $\mathbb{E}(Z) = 1 \times \mathbb{P}(Z=1) + 0 \times \mathbb{P}(Z=0)$

$= 1 \times \mathbb{P}(X \geq \delta) = \mathbb{P}(X \geq \delta)$

So $\mathbb{E}(X) \geq \delta \mathbb{P}(X \geq \delta)$ and

$$\mathbb{P}(X \geq \delta) \leq \frac{\mathbb{E}(X)}{\delta} \quad \boxtimes$$

Week 6

1

Example. A person wants to pass a driving test. On average, one needs 2.5 attempts. Prove that the event that one will need 10 or more attempt is at most $\frac{1}{4}$.

Solution. X = the number of attempts. We know that $\mathbb{E}(X) = 2.5$. Hence

$$\mathbb{P}(X \geq 10) \leq \frac{\mathbb{E}(X)}{10} = \frac{2.5}{10} = \frac{1}{4}. \quad \boxtimes$$

The Chebyshev inequality

__Theorem.__ Let X be a r.v. with $\mathbb{E}(X)=\mu$, $Var(X)=\sigma^2$. Then for any $\varepsilon>0$

$$\mathbb{P}(|X-\mu|\geq\varepsilon)\leq\frac{Var(X)}{\varepsilon^2}\equiv\frac{\sigma^2}{\varepsilon^2}.$$

__Proof.__ Set $Y=(X-\mu)^2$. Then

$$\mathbb{P}(|X-\mu|\geq\varepsilon)=\mathbb{P}((X-\mu)^2\geq\varepsilon^2)$$

$$=\mathbb{P}(Y\geq\varepsilon^2).$$

Since $Y\geq 0$, we can use the Markov inequality:

$$\mathbb{P}(Y\geq\varepsilon^2)\leq\frac{\mathbb{E}(Y)}{\varepsilon^2}=\frac{\mathbb{E}((X-\mu)^2)}{\varepsilon^2}=\frac{Var(X)}{\varepsilon^2}.$$

$\boxtimes$

__Remark.__ We use one of the equivalent definitions of variance: $Var(X)=\mathbb{E}((X-\mu)^2)$, where $\mu=\mathbb{E}(X)$.

__Example.__ Consider a sequence of 150 independent trials. The pr-ty of success in each trial is $p=0.4$.

Etimate the pr-ty of 80 successes in 150 trials.

<u>Solution</u>. X = the number of successes

Then $X \sim \text{Binomial}(150, 0.4)$.

We know that $\mathbb{E}(X) = np = 150 \cdot 0.4 = 60$,

$$Var(X) = npq = 150 \cdot 0.4 \cdot 0.6 = 60 \cdot 0.6 = 36.$$

Then $\mathbb{P}(X \geq 80) = \mathbb{P}(\underbrace{X - 60}_{X-\mu} \geq 20) \leq$

$$\mathbb{P}(|X - 60| \geq 20) \leq \frac{Var(X)}{20^2} = \frac{36}{20^2} = \frac{9}{5 \cdot 20} = 0.09.$$

<u>The Law of Large Numbers (LLN)</u>

We need the following

<u>Lemma</u>. If $X_1, X_2, \dots, X_n$ are independent r.v.'s, then

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i)$$

<u>Remark</u>. $\mathbb{E}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathbb{E}(X_i)$ — this does not require the X_i's to be independent.

<u>Theorem</u>. [The LLN]. Let $X_1, X_2, \ldots, X_n$ be a sequence of independent r.v.'s with $\mathbb{E}(X_i) = \mu$, $Var(X_i) = \sigma^2$. Set

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

Then for any $\varepsilon > 0$

$$P(|Y_n - \mu| \leq \varepsilon) \to 1 \text{ as } n \to \infty.$$

<u>Proof</u>. $\mathbb{E}(Y_n) = \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\mathbb{E}\left(\sum_{i=1}^{n} X_i\right)$

$$= \frac{1}{n}\sum_{i=1}^{n} \underbrace{\mathbb{E}(X_i)}_{=\mu} = \frac{1}{n}\sum_{i=1}^{n} \mu = \mu.$$

$$Var(Y_n) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2} Var\left(\sum_{i=1}^{n} X_i\right)$$

By the above Lemma,

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \underbrace{Var(X_i)}_{\sigma^2} = n\sigma^2$$

Hence

$$Var(Y_n) = \frac{1}{n^2} \times n\sigma^2 = \frac{\sigma^2}{n}.$$

Now,

$$P(|Y_n - \mu| \geq \varepsilon) = P(|Y_n - \mathbb{E}(Y_n)| \geq \varepsilon)$$

$$\leq \frac{Var(Y_n)}{\varepsilon^2} \quad \text{(by the Chebyshev inequality)}.$$

so $\mathbb{P}(|Y_n - \mu| > \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2} \xrightarrow[n\to\infty]{} 0$

Hence $\mathbb{P}(|Y_n - \mu| \leq \varepsilon) =$

$1 - \underbrace{\mathbb{P}(|Y_n - \mu| > \varepsilon)}_{\to 0} \to 1$ as $n \to \infty$ $\boxed{}$

The Central Limit Theorem (CLT)

Theorem. Let $X_1, X_2, \dots X_n$ be independent identically distributed r.v.s with $\mathbb{E}(X_i) = \mu$, $Var(X_i) = \sigma^2$. Then

$$\mathbb{P}\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\,\sigma} < x\right) \xrightarrow[n\to\infty]{} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt.$$

Remarks 1) Denote $Z_n = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\,\sigma}$.

The CLT tells us that Z_n is approximately normal $(0,1)$:

$$Z_n \approx \sim N(0,1).$$

In applications, we suppose that $Z_n \sim N(0,1)$

2) $Y_n = \frac{1}{n} \sum_{i=1}^{n} X_i - \mu = \frac{1}{n}\left(\sum_{i=1}^{n} X_i - n\mu\right)$

so $Z_n = \frac{\sqrt{n}}{\sigma} Y_n$.

Tutorial 5, CW5.

$$\begin{cases} Y_1 = X_1/X_2 \\ Y_2 = X_2 \end{cases} \Longrightarrow \begin{cases} X_1 = Y_1 Y_2 . \\ X_2 = Y_2 \end{cases} \qquad x_1$$

$$J = \begin{vmatrix} y_2 & y_1 \\ 0 & 1 \end{vmatrix} = y_2 .$$

$$f_{Y_1,Y_2}(y_1, y_2) = f_{X_1 X_2}(y_1 y_2, y_2) \cdot y_2 = 8 y_1 y_2^3 .$$

$$0 < x_1 < x_2 < 1 \iff$$

$$0 < y_1 y_2 < y_2 < 1 .$$

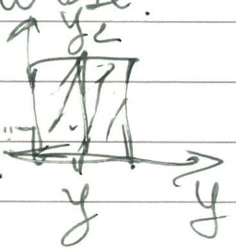
so $y_1 > 0,\ y_2 > 0;\quad y_2 < 1.$

$$y_1 y_2 < y_2 \Rightarrow y_1 < 1 .$$

so:

$$f_{Y_1 Y_2}(y_1, y_2) = \begin{cases} 8 y_1 y_2^3 & \text{if } 0 < y_1 < 1,\ 0 < y_2 < 1 . \\ 0 & \text{otherwise.} \end{cases}$$

$$f_{Y_1}(y_1) = \int_0^1 f_{Y_1 Y_2}(y_1, y_2)\, dy_2 = 2y_1 .$$



$$f_{Y_1}(y_1) = c y_1 = 2 y_1 .$$

For other problems, see typeset solutions.

Tutorial 6. CW6.

Facts used in CW5,6.

$X \geq 0, \quad \mathbb{P}(X \geq \delta) \leq \frac{\mathbb{E}(X)}{\delta}$ Markov's,

$\mathbb{P}(|X - \mu| \geq \varepsilon) \leq \frac{Var(X)}{\varepsilon^2}$ Tchebyshev's inequalities.

$\frac{1}{n} \sum_{i=1}^{n} X_i \to \mu$ LLN

CLT $S_n \stackrel{def}{=} \sum X_i \Longrightarrow \mathbb{P}\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\,\sigma} < x\right) \approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt = \Phi(x)$

CW6.

1) $Var\left(\sum_{i=1}^{n} X_i\right) = \mathbb{E}\left(\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \mu_i\right)^2$

$= \mathbb{E}\left(\sum_{i=1}^{n} (X_i - \mu_i)\right)^2$

$= \mathbb{E}\left(\sum_{i=1}^{n} (X_i - \mu_i)^2 + 2 \sum_{1 \leq i < j \leq n} (X_i - \mu_i)(X_j - \mu_j)\right)$

$= \sum_{i=1}^{n} \underbrace{\mathbb{E}(X_i - \mu_i)^2}_{Var(X_i)} + 2 \sum_{1 \leq i < j \leq n} \underbrace{\mathbb{E}(X_i - \mu_i)(X_j - \mu_j)}_{cov(X_i, X_j) = 0}$

$= \sum_{i=1}^{n} Var(X_i).$

2) $\mathbb{P}\left(\left|\frac{1}{n} \sum_{i=1}^{n} X_i - \mu\right| < 0.1\sigma\right) \geq 0.95 \Longleftrightarrow$

$\mathbb{P}(|\overline{X}_n - \mu| \geq 0.1\sigma) \leq \frac{\sigma^2}{n(0.1\sigma)^2} = \frac{100}{n} \leq 0.05$

$Var(\overline{X}_n) = \frac{\sigma^2}{n},$

$\Longrightarrow n \geq \frac{100}{0.05} = 2000.$

4) X_i = the hight of i-th person.

$$\frac{1}{100}\sum_{i=1}^{100} X_i \geq 163. \qquad \mathbb{E}(X_i) = 160, \; Var(X_i) = 15^2$$

$$\sum_{i=1}^{100} X_i \geq 100 \times 163$$

$$Z = \underbrace{\frac{\sum_{1}^{100} X_i - 100 \times 160}{10 \times 15}}_{\sim N(0,1)} \geq \frac{100 \times 163 - 100 \times 160}{10 \times 15} = \frac{300}{150} = 2$$

$$\mathbb{P}(Z \geq 2) = 1 - \Phi(2) = 0{,}023.$$

5) $U = X_1 + \dots + X_{1000}$ $\qquad E(X_i) = 0.45 - 0.55 = -0.1.$

$$Var(X_i) = \underset{\mathbb{E}(X_i^2)}{\underset{\parallel}{1}} - 0{,}1^2 = 0{,}99$$

$$\mathbb{P}(U < -20) =$$

$$\mathbb{P}\left(\frac{U + 1000 \times 0.1}{\sqrt{1000 \times 0{,}99}} < \frac{-20 + 100}{\sqrt{990}}\right)$$

$$= \mathbb{P}\left(\frac{U + 100}{\sqrt{990}} < 2.54\right) = \Phi(2{,}54) = 0{,}9943$$

$$\mathbb{P}(\text{Go bankrupt}) \geq 0.9943$$

Comments concerning exam

Understand that:

1) Example $\mathbb{P}(Y^2 > X > 0) = \int_0^\infty \left(\int_{\sqrt{x}}^\infty f_{X,Y}(x,y)\,dy \right) dx$

$$= \int_0^\infty \left(\int_0^{y^2} f_{XY}(x,y)\,dx \right) dy$$

Know what is $f_{X|Y=y}$, $f_{Y|X=x}$.

2) Understand: $\mathbb{E}(g(X) | Y=y) = \int_{-\infty}^{\infty} g(x) f_{X|Y=y}(x)\,dx$

Be able to prove <u>all</u> statements we proved in lectures. Examples:

a) $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$, $f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$

(b) X, Y are independent $\Leftrightarrow f_{X,Y}(x,y) = h(x)g(y)$.

(c) $X_1, \ldots, X_n$ are normal r.v.'s, $X_i \sim N(\mu_i, \sigma_i^2)$.

Then $\sum_{i=1}^n X_i \sim N\left(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2\right)$.

(d) $X_i \sim Exp(\lambda)$, What is the pdf of $\sum_{i=1}^n X_i$

(e) Inequalities: Markov, Chebyshev.

(f) LLN. Etc.

<u>In other words:</u> Know <u>all</u> proofs!