§0. prerequisite

I. common itg / drvative

1. $\int f(x) g'(x) dx = F(x)\cdot g(x) - \int F(x)\, g'(x) dx$.   2. $\begin{cases}\int e^{kx}dx = \frac{1}{k}e^{kx} \\ \int a^{kx}dx = \frac{1}{k}\frac{1}{\ln a}a^{kx}\end{cases}$ 3. $\int x^n dx = \frac{1}{n+1}x^{n+1}$ 4. $\int \frac{1}{x}dx = \ln|x|$

1. $(\frac{1}{x})' = -\frac{1}{x^2}$, $(\frac{1}{\sqrt{x}})' = \frac{1}{2\sqrt{x}}$, $(\log_a x)' = \frac{1}{x\ln a}$, $(a^x)' = \ln a \cdot a^x$, $(a^u)' = u'(a^u)'$ 2. $(fg)' = f'g + fg'$ 3. $(\frac{f}{g})' = \frac{f'g - fg'}{g^2}$

II. central lim thm Δ: mind σ or σ^2

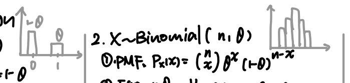
1. def (many IIDs sum is normal distribution)
$X_1, X_2, \dots X_n$, idpd with same PMF/PDF (iid), $X = X_1+X_2+\dots+X_n$, $E(X_i) = \mu$, $Var(X_i) = \sigma^2$
then for every t: $\lim_{n\to\infty} P(X \le \mu + t\sigma) = \Phi(t)$, $(\Phi(t)$: CDF of $N(0,1))$
→ or: $\lim_{n\to\infty}(\frac{X-\mu}{\sigma} \le t) = \Phi(t)$ or: $X \sim N(\mu, \sigma)$ → note ①: n as large as 30 is OK to use this

III. commonly used pbb formula:

①. $P(A) = \frac{|A|}{|\Omega|}$ ②. $P(A|F) = \frac{P(A\cap F)}{P(F)} = \frac{|A\cap F|}{|F|}$ ③. $P(A\cap B) = P(A|B)\cdot P(B)$
④. $P(A) = P(A\cap B) + P(A\cap B^c) = P(A|B)\cdot P(B) + P(A|B^c)\cdot P(B^c)$ → $P(A) = \sum_B P(A,B)$
⑤. $P(A|B) = \frac{P(B|A)\cdot P(A)}{P(B)}$ → $P(a|b,c) = \frac{P(b|a,c)\cdot P(a|c)}{P(b|c)}$ see cas universal cdt; a, b, c can be a set of events
⑥. A·B idpd ⇔ $P(A\cap B) = P(A)\cdot P(B)$ ⇔ $P(A) = P(A|B)$; often idpd add universal condition.

☆ write PMF in table

IV. diff distribution



A.1. $X\sim$ Bernoulli(θ)
①. PMF: $P\begin{cases}P(X=0) = 1-\theta \\ P(X=1) = \theta\end{cases}$
②. $E(X) = \theta$, $Var(X) = \theta(1-\theta)$
③. $\theta \in [0,1]$

2. $X\sim$ Binomial(n, θ)
①. PMF: $P_X(x) = \binom{n}{x}\theta^x(1-\theta)^{n-x}$
②. $E(X) = n\theta$, $Var(X) = n\theta(1-\theta)$
③. $\theta \in [0,1]$

3. $X\sim$ Geometric(θ)
①. PMF: $P_X(x) = (1-\theta)^{x-1}\theta$
②. $E(X) = \frac{1}{\theta}$, $Var(X) = \frac{1-\theta}{\theta^2}$
③. $\theta \in [0,1]$

4. $X\sim$ Poisson(θ)
①. PMF: $P_X(x) = e^{-\theta}\frac{\theta^x}{x!}$
②. $E(X) = \theta$, $Var(X) = \theta$
③. $\theta \in (0,\infty)$

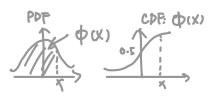
5. $X\sim$ Uniform(a,b)
①. PDF: $f(t) = \begin{cases}\frac{1}{b-a} & (a\le t\le b) \\ 0 & (\text{otherwise})\end{cases}$
②. CDF: $F(t) = P(T\le t) = \begin{cases}0 & (t\le a) \\ \frac{1}{b-a}t - \frac{a}{b-a} & (a<t<b) \\ 1 & (t\ge b)\end{cases}$
③. $E(X) = \frac{a+b}{2}$, $Var(X) = \frac{(b-a)^2}{12}$

6. $X\sim$ Exponential(θ)
①. PDF: $f_X(x) = \begin{cases}\theta e^{-\theta x} & (x\ge 0) \\ 0 & (\text{otherwise})\end{cases}$
②. CDF: $F_X(x) = \begin{cases}1-e^{-\theta x} & (x>0) \\ 0 & (\text{otherwise})\end{cases}$
③. $E(X) = \frac{1}{\theta}$, $Var(X) = \frac{1}{\theta^2}$
④ $\theta \in (0,\infty)$

7. $X\sim N(\mu, \sigma^2)$
①. PDF: $f(x) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
②. $E(X) = \mu$, $Var(X) = \sigma^2$
③. argmax $= \mu$.

$Z = \frac{X-\mu}{\sigma}$ ⇄ $X = \sigma Z + \mu$

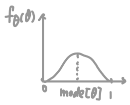
8. $X\sim N(0,1)$
①. PDF: $f_X(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$
②. CDF: $\Phi(x)$



④: other ppt: <a>. idpd $X\sim N(\mu_X, \sigma_X^2)$, $Y\sim N(\mu_Y, \sigma_Y^2)$ ⇒ $X+Y \sim N(\mu_X+\mu_Y, \sigma_X^2+\sigma_Y^2)$
idpd $X_i \sim N(\mu_i, \sigma_i^2)$ ⇒ $\sum_{i=1}^n a_iX_i \sim N(\sum_{i=1}^n a_i\mu_i, \sum_{i=1}^n (a_i\sigma_i)^2)$

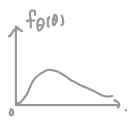
B. common priors: Beta(α,β) and Gamma(α,β)
0. prqs: $\Gamma(t) = \int_0^\infty x^{t-1}e^{-x}dx$; when t: +ve int, $\Gamma(t) = (t-1)!$

1. $\theta \sim$ Beta(α,β)
①. $f_\theta(\theta) = \begin{cases}\frac{1}{B(\alpha,\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1} & (0<\theta<1) \\ 0 & (\text{otherwise})\end{cases}$, $B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$



②. mode[θ] $= \frac{\alpha-1}{\alpha+\beta-2}$
③. notes, <1>. Uniform(0,1) ~ Beta(1,1)
<2>. $\int_0^1 \theta^{\alpha-1}(1-\theta)^{\beta-1}d\theta = B(\alpha,\beta)$

2. $\theta \sim$ Gamma(α,β)
①. $f_\theta(\theta) = \begin{cases}\frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta} & (\theta>0) \\ 0 & (\text{otherwise})\end{cases}$



②. notes,
<1>. $\int_0^\infty \theta^{\alpha-1}e^{-\beta\theta}d\theta = \frac{\Gamma(\alpha)}{\beta^\alpha}$
(<2>) not proven mode: when $\alpha\ge 1$: mode $= \frac{\alpha-1}{\beta}$; $\alpha<1$: decreases, no mode
<3>. common: $\int\theta^a e^{-b\theta}d\theta = -\frac{e^{-b\theta}}{b}\sum_{k=0}^a\frac{\theta^{a-k}(a)_k}{b^k}$; $(a)_k = \overbrace{a(a-1)\cdots(a-k+1)}^{k}$, $a\in\mathbb{N}$

§1. Bayesian Statistics

I. Bayesian statistics inference prior $\xrightarrow{\text{experiments}}$ posterior

1. single experiment
① draw graph / curve of $X|\theta$ and θ
② find all p/f $\theta|X(\theta|x)$, with term $Z(X)$, ↙ mind range of θ, handle "small equations" separately; this cannot need θe?...
$P_{\theta|X}(\theta|x) = \frac{P_{X|\theta}(x|\theta)\cdot P_\theta(\theta)}{Z(X)}$ ↙ p stands for p/f
③ find $Z(X)$: $\begin{cases}\text{if } \theta \text{ discrete: } \sum_\theta P_{\theta|X} = 1 \Rightarrow Z(X) = \sum \text{numerator} \\ \text{if } \theta \text{ continuous: } \int f_{\theta|X}d\theta = 1 \Rightarrow \frac{1}{Z(X)} = \frac{1}{\int \text{numerator}\, d\theta}\end{cases}$

2. multiple experiment. eg: X_1, X_2
<1>. already got $P_{\theta|X_1}$: $P_{\theta|X_2X_1}(\theta|x_2,x_1) = \frac{P_{X_2|\theta,X_1}(x_2|\theta,x_1)\cdot P_{\theta|X_1}(\theta|x_1)}{Z(X_2,X_1)} = \frac{P_{X_2|\theta}(x_2|\theta)\cdot P_{\theta|X_1}(\theta|x_1)}{Z(X_2,X_1)}$
<2>. haven't got $P_{\theta|X_1}$: $P_{\theta|X_2,X_1}(\theta|x_2,x_1) = \frac{P_{X_2,X_1|\theta}(x_2,x_1|\theta)\cdot P_\theta(\theta)}{Z(X_2,X_1)} = \frac{P_{X_1|\theta}(x_1|\theta)P_{X_2|\theta}(x_2|\theta)\cdot P_\theta(\theta)}{Z(X_2,X_1)}$

3. notes
<1>. when ask: find posterior PDF of ... can write: PDF is Beta(9,8)

II. conjugate expression → if prior and posterior same dist. family, called cjgt dist.

1. thms. to update param for experiment's distr.
① $\theta\sim$ Beta(α,β) $\xrightarrow{\text{many Bernoulli}(\theta)}$ Beta(α+success, β+fail)
(②) $\theta\sim$ Beta(α,β) $\xrightarrow{n * \text{Geometric}(\theta)}$ Beta($\alpha+n$, $\beta+\sum x_i - n$) ← not proven
③ $\theta\sim$ Gamma(α,β) $\xrightarrow{n * \text{Poisson}(\theta)}$ Gamma($\alpha+\sum x_i$, $\beta+n$)
④. $\theta\sim$ Gamma(α,β) $\xrightarrow{n * \text{Exponential}(\theta)}$ Gamma($\alpha+n$, $\beta+\sum x_i$)
⑤. $\mu\sim N(\mu_0, \sigma_0^2)$ $\xrightarrow{n * \text{Normal}(\mu, \sigma^2)}$ $N(\mu', \sigma'^2)$: where in samples μ is unknown and σ^2 known
$\mu' = \frac{\sigma^2\mu_0 + \sigma_0^2\sum X_i}{\sigma^2 + n\sigma_0^2}$, $\sigma'^2 = \frac{\sigma^2\sigma_0^2}{\sigma^2+n\sigma_0^2}$

III. prediction, estimation, hypo testing

1. prediction: $P_{X_2|X_1}(x_2|x_1)$
① find posterior ← with θe? ②: write any p/f, $\int/\Sigma$ on draft as: $P(X_2=x_2|X_1=x_1) = \int_{-\infty}^{\infty} P(X_2=x_2|\theta=\theta)\cdot P(\theta=\theta|X_1=x_1)d\theta$
③ pick p/f, $\int/\Sigma$ cal inner itg (Δ mind range), ④ find range of $\int/\Sigma$. write down with intermdt steps ⑤. calculate outmost itg.

2. point estimation: $\hat\theta_{MAP}$ = arg max$_\theta$ $f_{\theta|X}(\theta|x)$
①. general: ¹find posterior ²find extreme value at boundary or $f'=0$ ②: common dist: ppt. eg: Beta, Gamma's mode

3. hypo testing: best θ in candidates: in problem: it is still called: "estimate based on MAP rule": max for candidate θ
①. general: $\begin{cases}\text{best } \theta: \text{ find all } f_{\theta|X} \text{ and find } \theta \text{ in candidates that make } f_{\theta|X} \text{ max} \\ \text{error: the pbb that wrong estimate based on MAP rule: find } P(\hat\theta_{MAP}\ne\theta) \text{ by sum rule}\end{cases}$
②: common, binary hypo:

best θ: $\frac{f_{\theta|X}(1|x)}{f_{\theta|X}(0|x)} = \frac{f_{X|\theta}(x|1)\cdot f_\theta(1)}{f_{X|\theta}(x|0)\cdot f_\theta(0)}$ $\begin{cases}>1: \hat\theta=1 \\ \le 1: \hat\theta=0\end{cases}$

error: $P(\hat\theta\ne\theta) = \begin{cases}P(\hat\theta=0,\theta=1) + P(\hat\theta=1,\theta=0) = P(\hat\theta=0|\theta=1)\cdot P(\theta=1) + P(\hat\theta=1|\theta=0)\cdot P(\theta=0) & \text{not given } X\text{, unknown } \hat\theta \\ = P(\theta\ne\hat\theta|X) = \frac{P(X|\theta=\hat\theta)\cdot P(\theta=\hat\theta)}{Z(X)} & \text{: by Bayes statistics inference, already cal} \quad \text{given } X, \hat\theta \text{ already calculated}\end{cases}$
MAP

note: in this course, $\begin{cases}sd: \sigma = \sqrt{var} = \sqrt{\frac{1}{n}\sum_{i=1}^n(x_i-\mu)^2} & (\text{for real distr}) \\ se: \sigma_{\bar X} = \sigma/\sqrt{n} & (\text{for actual smpl})\end{cases}$

§2. sample statistics

I. defs. population, random sample (of size n) $(X_1, X_2, \dots)$, sample, sampling, sampling stats, spling distr. ← spl of idpd RV

II. SS measures
1. sample num n. 2. sample proportion $\hat p = \frac{\text{num of target}}{n}$ 3. sample sum $X = X_1+\dots+X_n$
4. sample mean $\bar X$.
①. def. $\bar X = \frac{X_1+\dots+X_n}{n}$ ②. opposite: actual / ppletion mean: $\mu = E(X) = \sum_x xp(x)$ (here x is ele in pplt, not spl sum) ↳ if pplt follow distr. just use...
③. ppt1: $E(\bar X) = \mu$, $\bar X$ is unbiased estimator of μ; (not proven): $Var(\bar X) = \frac{\sigma^2}{n}$
④. ppt2. <1>. ccls, $n\to\infty$, $\bar X \sim N(\mu, (\frac{\sigma}{\sqrt n})^2)$, $\mu_{\bar X} = \mu$, $\sigma_{\bar X} = \frac{\sigma}{\sqrt n}$ (by CLT)
<2>. ROT: if population normal, $\bar X$ must be normal; if population not normal, $n\ge 30$: $\bar X$ consider normal.
<3>: other expression: $P(|\bar X - \mu| \ge \varepsilon) \le \delta$, where ε: spling error, δ: confidence error

5. sample variance s^2, sample std dev (std error) s:
①. def. $s^2 = \frac{\sum(X_i-\bar X)^2}{n} = \frac{\sum X_i^2}{n} - \bar X^2$; $s = \sqrt{s^2}$
②. opposite: pplt/actual var σ^2, dev σ; $\sigma^2 = E[(x-\mu)^2] = \sum_x (x-\mu)^2 P(x) = E(X^2) - E(X)^2$ (where x ...)
③. ppt $E(s^2) = \frac{n-1}{n}\sigma^2$, biased estimator of σ^2
④. unbiased spl var: $\frac{n}{n-1}s^2 = \frac{\sum(X_i-\bar X)^2}{n-1} = \frac{\sum_{i=1}^n X_i^2}{n-1} - \frac{n}{n-1}\bar X^2$
$E(\frac{n}{n-1}s^2) = \sigma^2$
☆ whenever say: estimate var from sample: unbiased var $\frac{n}{n-1}s^2$

III. estimator
1. desc. if sth can estimate (no matter good/bad) sth, call it estimator. (eg. $\bar X$ is estimator of μ)
2. features:
unbiased: $E[\hat\theta_n] = \theta$; asymptotically unbiased: $\lim_{n\to\infty}E[\hat\theta_n] = \theta$
consistent: $\hat\theta_n$ converges to θ in pbb. $\lim_{\varepsilon\to 0}\lim_{n\to\infty} P(|\hat\theta_n - \theta| \ge \varepsilon) = 0$

§3 classical stats: θ considered as deterministic quantity

I. point estimation
1. estimator $\hat\theta_n = g(X_1, \dots X_n)$: is a fn, input experiment data, output est. θ
2. maximum likelihood estimation: $\hat\theta_n$ = arg max$_\theta$ $f_X(x_1,\dots,x_n;\theta)$: extreme maxima
①. ☆ discrete: compute $f_X(x_1,\dots x_n;\theta)$ for each θ, $\hat\theta = \theta_{MLE}$ = argmax
②. ☆ continuous and differentiable
w1:
A. first write down θ_{MLE} = arg max$_\theta$ $f_X(x_1,\dots x_n;\theta)$ = argmax$_\theta$$\{\ln[f_X(x_1|\theta)] + \dots + \ln[f_X(x_n|\theta)]\}$
B. cal with x_i: $\frac{\partial}{\partial\theta}\ln[f_X(x_i;\theta)]$
C. solve: $\frac{\partial}{\partial\theta}\ln[f_X(x_1;\theta)] + \dots + \frac{\partial}{\partial\theta}\ln[f_X(x_n;\theta)] = 0$: get extreme point
D. compare end point and all extreme pts in C, find arg max$_\theta$
w2:
A. first write down θ_{MLE} = arg max$_\theta$ $f_X(x_1,\dots x_n;\theta)$
B. write down $f_X(x_1,\dots x_n;\theta) = f_X(x_1;\theta)\cdots f_X(x_n;\theta) = \cdots$
C. solve: $\frac{\partial}{\partial\theta}[f_X(x_1;\theta)\cdots f_X(x_n;\theta)] = 0$
D. compare end point and all extreme pts in C, find arg max$_\theta$

③. ☆. multivar cnt. $\hat\theta_1, \dots \hat\theta_m$ = argmax$_{\theta_1\dots\theta_m}$ $f_X(x_1,\dots x_n;\theta_1,\dots\theta_m)$
☆. $\frac{\partial}{\partial\theta_1}\ln[f_X(x_1;\theta)] + \dots + \frac{\partial}{\partial\theta_1}\ln[f_X(x_n;\theta)] = 0$; ... ; $\frac{\partial}{\partial\theta_m}\ln[f_X(x_1;\theta)] + \dots + \frac{\partial}{\partial\theta_m}\ln[f_X(x_n;\theta)] = 0$;
other same with ②

④. common cases:
<1>. $X_1, \dots X_n$ idpd from Uniform(0,θ) ⇒ $\theta_{MLE} = \max\{x_1,\dots x_n\}$
<2>. X ~ Bernoulli(θ), k succ, n-k fail ⇒ $\theta_{MLE} = \frac{k}{n}$
<3>. $X\sim N(\mu, \sigma^2)$: μ_{MLE}, σ_{MLE} same with spl mean, spl var

II. ~~prediction~~: confidence interval of ppl mean μ

1. defs: ①. $[\hat\theta_n^-, \hat\theta_n^+]$ is called a $(1-\alpha)$ - confidence interval if $P(\hat\theta_n^- \le \theta \le \hat\theta_n^+) \ge 1-\alpha$
②. $\hat\theta_n^-, \hat\theta_n^+$: lower, higher confidence limit; $\hat\theta_n^+ - \hat\theta_n^-$: width; α: cfd param, $1-\alpha$: cfd. level (α: significant level)

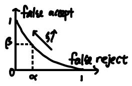
2. ☆: (narrowest) cfd itv of IID $X_1, \dots X_n$, ppl mean μ
①. σ^2 known: $n\ge 30$ non-normal; n ~ normal: $\bar X \pm z_{\frac{\alpha}{2}}\cdot(\frac{\sigma}{\sqrt n})$
②. σ^2 unknown: $n\ge 30$ non-normal; $n\ge 30$ normal: $\bar X \pm z_{\frac{\alpha}{2}}\cdot(\frac{s}{\sqrt n})$
③ σ^2 unknown: $n<30$ normal: $\bar X \pm t_{\frac{\alpha}{2}}\cdot(\frac{s}{\sqrt n})$; t is of $n-1$ deg. of freedom
→ note: <1>. find $z_{\frac{\alpha}{2}}$, N(0,1): $-z_{\frac{\alpha}{2}}$ $z_{\frac{\alpha}{2}}$, ①. find $1-\frac{\alpha}{2}$ in z-value table ② crspd value
find $t_{\frac{\alpha}{2}}$, t-distr: $-t_{\frac{\alpha}{2}}$ $t_{\frac{\alpha}{2}}$ ①. find crsp DoF ②. find $\frac{\alpha}{2}$ in t-value table ③. crsp value
<3>. S is unbiased. ☆ Bernolli, special case. use: $S = \sqrt{\hat p(1-\hat p)} = \sqrt{\bar x(1-\bar x)}$;
<4>. t - dist:
[1]. def. $t(n) \sim T = \frac{Z}{\sqrt{\frac{Y}{n}}} = \frac{N(0,1)}{\sqrt{\frac{\chi^2(n)}{n}}}$; $\chi^2(n) = (x_1^2+\dots+x_n^2)$, x_i idpd N(0,1)
[2]. ppt: n↑, more like normal [3]. table t_α: α → shaded value
[4]. thm. if $X_1, \dots X_n$: idpd $N(\mu, \sigma^2)$, $T = \frac{\bar X - \mu}{\frac{s}{\sqrt n}} \sim t(n-1)$

Ⅲ. binary (simple) hypo testing:

△. 2 ways for s in multi-ppl △ s unbiased; only Bernoulli special case
☆: can round to 2 dec in intermidiate steps

1. def ① general hypo testing: $\theta_1 \cdots \theta_m$, select optimal θ^*
② binary hypo testing: θ is 0 or 1
<1>. terminology: [1]: decision rule: the way/fn to decide [2]: H_0: null hypo, $\theta=0$; H_1: alternate hypo, $\theta=1$
<2>. R: rejection region: H_0 rejected (H_1); R^c: acceptance region: H_0 accepted (H_0); find dec rule: choose R
<3>. type I error (false rejection): decide H_1 (reject H_0) $| H_0$; $\alpha(R) = P(x \in R; H_0) = P(L > \xi; H_0)$
type II error (false acceptance): decide $H_0 | H_1$; $\beta(R) = P(x \notin R; H_1) = P(L \le \xi; H_1)$

2. general dcs rule:
①: likelihood ratio (rule):
$$L(x_1, x_2, \dots x_n) = \frac{f_x(x_1, \dots x_n; H_1)}{f_x(x_1, \dots, x_n; H_0)} \; ; \; \begin{cases} \text{if } L(x_1,\dots x_n) > \xi : H_1 \text{ true (reject)} \\ \text{if } L(x_1, \dots x_n) \le \xi : H_0 \text{ true (accept)} \end{cases}, \; \xi > 0: \text{critical value}$$
②: special case:
<1>: MLE-based: $\xi = 1$; <2>. MAP-based: $\xi = \frac{P(H_0)}{P(H_1)}$
③. trade-off of ξ



3. llh ratio test: design decision rule
①: process (assume known dist. of X under H_0 and H_1)
[0] find what's X, H_0, H_1, find $f_{x;H_1}$, $f_{x;H_0}$, L
[1]: decide α (tpc: 0.01, 0.05, 0.1)
[2]: $\alpha = P(L(x) > \xi; H_0) = P(x > \cdots \xi; H_0) \Rightarrow$ find $\xi \Rightarrow \begin{cases} H_0, x \in \cdots \\ H_1, x \in \cdots \end{cases}$
[3]: find $\beta = P(x \in \cdots (H_0); H_1)$ ($= P(L(x) \le \xi; H_1) = P(x > \cdots \xi; H_1)$)   †: note: use some intermediate values to avoid error from rounding
②: neyman-pearson lemma: when α fixed, LRT finds the best β.

Ⅳ: binary composite hypo of ppl mean μ (1 ppl)

1. def. composite hypo: param not a value eg. $H: \mu \ne 30$, $\mu > 30$
2. prob desc: binary composite hypo about μ: given H_0, H_1 and significance level α (just false_rej pbb) either rej or acc H_0
3. critical value approach: assume H_0 true, find R with false-rej pbb $= \alpha$; estimate $\bar{x}$, if $\bar{x}$ in R: ✓ else ✗
<a>. 2-sided: $H_0: \mu = \mu_0$; $H_1: \mu \ne \mu_0$ (like finding cfd interval)
①: $n \ge 30$: $\begin{cases} \sigma^2 \text{ known: if } \frac{|\bar{x} - \mu_0|}{\sigma/\sqrt{n}} > z_{\frac{\alpha}{2}}, \text{ reject, else accept} \\ \sigma^2 \text{ unknown: change } \sigma \text{ to } S \end{cases}$
②: $n < 30, x_i \sim N(\mu, \sigma^2)$ $\begin{cases} \sigma^2 \text{ known: if } \frac{|\bar{x} - \mu_0|}{\sigma/\sqrt{n}} > z_{\frac{\alpha}{2}}: \text{reject, else don't} \\ \sigma^2 \text{ unknown: if } \frac{|\bar{x} - \mu_0|}{S/\sqrt{n}} > t_{\frac{\alpha}{2}}: (\text{D.F}: n-1): \text{reject, else don't.} \end{cases}$
→ proof. H_0

: shaded are rejection region R
<b>. one sided test:
①: convert: $\begin{cases} \text{right-sided: } H_0: \mu \le \mu_0 \text{ vs. } H_1: \mu > \mu_0 \Rightarrow H_0: \mu = \mu_0 \text{ vs. } H_1: \mu > \mu_0 \\ \text{left-sided: } H_0: \mu \ge \mu_0 \text{ vs. } H_1: \mu < \mu_0 \Rightarrow H_0: \mu = \mu_0 \text{ vs. } H_1: \mu < \mu_0 \end{cases}$ proof: eg. $\alpha = P(\text{rej } H_0; H_0: \mu = \mu_0)$ $\ge P(\text{rej } H_0; H_0: \mu \le \mu_0)$
② sol: right-sided
$$\begin{cases} n \ge 30: \begin{cases} \sigma^2 \text{ known: if } \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} > z_\alpha : \text{rej else acc} \\ \sigma^2 \text{ unknown: if } \frac{\bar{x} - \mu_0}{S/\sqrt{n}} > z_\alpha : \text{rej else acc} \end{cases} \\ n < 30, x_i \sim N(\mu, \sigma^2) \begin{cases} \sigma \text{ known: if } \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} > z_\alpha : \text{rej else acc} \\ \sigma^2 \text{ unknown: if } \frac{\bar{x} - \mu_0}{S/\sqrt{n}} > t_\alpha \; (\text{D.F}: n-1): \text{rej else acc} \end{cases} \end{cases}$$
③ sol: left-sided
$$\begin{cases} n \ge 30: \begin{cases} \sigma^2 \text{ known: if } \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} < -z_\alpha : \text{rej else acc} \\ \sigma^2 \text{ unknown: if } \frac{\bar{x} - \mu_0}{S/\sqrt{n}} < -z_\alpha : \text{rej else acc} \end{cases} \\ n < 30, x_i \sim N(\mu, \sigma^2) \begin{cases} \sigma \text{ known: if } \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} < -z_\alpha : \text{rej else acc} \\ \sigma^2 \text{ unknown: if } \frac{\bar{x} - \mu_0}{S/\sqrt{n}} < -t_\alpha : (\text{D.F}: n-1): \text{rej else acc} \end{cases} \end{cases}$$
→ proof. H_0

: shaded are rejection region R; or left tailed

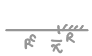
4. p-value: estm normalized $\bar{x}$; compute p-value, compare with α ($\bar{x}$: the mean of this stat; $\bar{X}$: variable: sample mean
→ def of p-value: given the specific stat $\bar{x}$, will reject H_0, smallest pbb of false rej error for which rej H_0
→. i.e. when $\bar{x}$ is the boundary, find the false-rej pbb, then compare with α

①: $n \ge 30$
$H_0: \mu = \mu_0$ vs. $H_1: \mu \ne \mu_0$ (two sided) $\begin{cases} (\sigma \text{ known}) \text{ p-value} = P\left(z > \left|\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}\right|\right) + P\left(z < -\left|\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}\right|\right) \\ (\sigma \text{ unknown}) \text{ p-value} \approx P\left(z > \left|\frac{\bar{x} - \mu_0}{S/\sqrt{n}}\right|\right) + P\left(z < -\left|\frac{\bar{x} - \mu_0}{S/\sqrt{n}}\right|\right) \end{cases}$
$H_0: \mu = \mu_0$ vs. $H_1: \mu > \mu_0$: (right sided) $\begin{cases} (\sigma \text{ known}) \text{ p-value} = P\left(z > \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}\right) \\ (\sigma \text{ unknown}) \text{ p-value} \approx P\left(z > \frac{\bar{x} - \mu_0}{S/\sqrt{n}}\right) \end{cases}$ $< \alpha$, rej; $\ge \alpha$, don't
$H_0: \mu = \mu_0$ vs. $H_1: \mu < \mu_0$: (left sided) $\begin{cases} (\sigma \text{ known}) \text{ p-value} = P\left(z < \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}\right) \\ (\sigma \text{ unknown}) \text{ p-value} \approx P\left(z < \frac{\bar{x} - \mu_0}{S/\sqrt{n}}\right) \end{cases}$

②: $n < 30$, $x_i \sim N(\mu, \sigma^2)$

H_0: $\mu = \mu_0$ vs. H_1: $\mu \neq \mu_0$ (two)
$$p\text{-value} \begin{cases} = P\left(Z > \left|\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}\right|\right) + P\left(Z < -\left|\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}\right|\right) & \sigma^2 \text{ is known} \\ \approx P\left(T > \left|\frac{\bar{x} - \mu_0}{s/\sqrt{n}}\right|\right) + P\left(T < -\left|\frac{\bar{x} - \mu_0}{s/\sqrt{n}}\right|\right) & \sigma^2 \text{ is unknown (D.F: n-1)} \end{cases}$$
H_0: $\mu = \mu_0$ vs. H_1: $\mu > \mu_0$ (right)
$$p\text{-value} \begin{cases} = P\left(Z > \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}\right) & \sigma^2 \text{ is known} \\ \approx P\left(T > \frac{\bar{x} - \mu_0}{s/\sqrt{n}}\right) & \sigma^2 \text{ is unknown (D.F: n-1)} \end{cases}$$
$< \alpha$, rej; $\ge \alpha$, don't
H_0: $\mu = \mu_0$ vs. H_1: $\mu < \mu_0$ (left)
$$p\text{-value} \begin{cases} = P\left(Z < \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}\right) & \sigma^2 \text{ is known} \\ \approx P\left(T < \frac{\bar{x} - \mu_0}{s/\sqrt{n}}\right) & \sigma^2 \text{ is unknown (D.F: n-1)} \end{cases}$$

proof:



p-value $= P(\bar{X} \in R; \mu = \mu_0)$

Ⅴ: binary composite hypo of ppl mean μ between 2 ppl.

1. prob desc. $X_1, \dots, X_{n_x}$ idpd with same μ_x, σ_x^2, $Y_1, \dots, Y_{n_y}$ idpd with same μ_y, σ_y^2, $\{X_i\}$'s, $\{Y_i\}$'s are also idpd.
2. Two-sided: $H_0: \mu_x = \mu_y$ vs. $H_1: \mu_x \ne \mu_y \longrightarrow H_0: \mu_x - \mu_y = 0$ vs. $H_1: \mu_x - \mu_y \ne 0$
① $n_x, n_y \ge 30$,
Critical val: $\begin{cases} \sigma_x^2, \sigma_y^2 \text{ known: if } \frac{|\bar{x} - \bar{y}|}{\sigma_D} > z_{\frac{\alpha}{2}}, \text{ rej} & \text{where } \sigma_D = \sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}} \\ \sigma_x^2, \sigma_y^2 \text{ unknown: if } \frac{|\bar{x} - \bar{y}|}{S_D} > z_{\frac{\alpha}{2}}, \text{ rej} & \text{where } S_D = \sqrt{\frac{S_x^2}{n_x} + \frac{S_y^2}{n_y}} \end{cases}$
p-val: $\begin{cases} \sigma_x^2, \sigma_y^2 \text{ known: p-val} = P\left(|z| > \left|\frac{\bar{x} - \bar{y}}{\sigma_D}\right|\right) < \alpha, \text{ rej} & \text{where } \sigma_D = \sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}} \\ \sigma_x^2, \sigma_y^2 \text{ unknown: p-val} = P\left(|z| > \left|\frac{\bar{x} - \bar{y}}{S_D}\right|\right) < \alpha, \text{ rej} & \text{where } S_D = \sqrt{\frac{S_x^2}{n_x} + \frac{S_y^2}{n_y}} \end{cases}$
②: $n_x, n_y < 30$, X_i, Y_i normal:
Critical val: $\begin{cases} \sigma_x^2, \sigma_y^2 \text{ known: if } \frac{|\bar{x} - \bar{y}|}{\sigma_D} > z_{\frac{\alpha}{2}}, \text{ rej; where } \sigma_D = \sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}} \\ \sigma_x^2, \sigma_y^2 \text{ unknown: if } \frac{|\bar{x} - \bar{y}|}{S_D\sqrt{1/n_x + 1/n_y}} > t_{\alpha/2} \; (\text{D.F}: n_x + n_y - 2): \text{, rej;} \end{cases}$
where pooled $S_D = \sqrt{\frac{(n_x - 1)S_x^2 + (n_y - 1)S_y^2}{n_x + n_y - 2}}$
p-val: $\begin{cases} \sigma_x^2, \sigma_y^2 \text{ known: p-value} = P\left(|z| > \frac{|\bar{x} - \bar{y}|}{\sigma_D}\right) < \alpha, \text{ rej, where } \sigma_D = \sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}} \\ \sigma_x^2, \sigma_y^2 \text{ unknown p-value} = P\left(|T| > \frac{|\bar{x} - \bar{y}|}{S_D\sqrt{1/n_x + 1/n_y}}\right) < \alpha, \text{ rej, (D.F}: n_x + n_y - 2) \end{cases}$
where pooled $S_D = \sqrt{\frac{(n_x - 1)S_x^2 + (n_y - 1)S_y^2}{n_x + n_y - 2}}$

proof: normal situ: $\bar{X} - \bar{Y} \sim N(\underbrace{\mu_x - \mu_y}_{0}, \frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y})$, $z = \frac{\bar{X} - \bar{Y}}{\sigma_D} \sim N(0,1)$; t-situ: $T = \frac{(\bar{X} - \bar{Y}) - (\mu_x - \mu_y)}{S_D\sqrt{1/n_x + 1/n_y}} \sim t(n_x + n_y - 2)$

3. left-sided: $H_0: \mu_x = \mu_y$ vs. $H_1: \mu_x > \mu_y \longrightarrow H_0: \mu_x - \mu_y = 0$ vs. $H_1: \mu_x - \mu_y > 0$
①. known σ; $n_x, n_y \ge 30$ or $n_x, n_y < 30$, both normal
Critical value: if $\frac{\bar{x} - \bar{y}}{\sigma_D} > z_\alpha$, rej; where $\sigma_D = \sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}$
p-value: p-val $= P\left(z > \frac{\bar{x} - \bar{y}}{\sigma_D}\right) < \alpha$, rej where $\sigma_D = \sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}$
②. unknown σ; $n_x, n_y \ge 30$
Critical value: if $\frac{\bar{x} - \bar{y}}{S_D} > z_\alpha$, rej; where $S_D = \sqrt{\frac{S_x^2}{n_x} + \frac{S_y^2}{n_y}}$
p-value: p-val $= P\left(z > \frac{\bar{x} - \bar{y}}{S_D}\right) < \alpha$, rej where $S_D = \sqrt{\frac{S_x^2}{n_x} + \frac{S_y^2}{n_y}}$
③ σ unknown; $n_x, n_y < 30$, both normal
Critical value: if $\frac{\bar{x} - \bar{y}}{S_D\sqrt{1/n_x + 1/n_y}} > t_\alpha$ (D.F: $n_x + n_y - 2$): , rej
where pooled $S_D = \sqrt{\frac{(n_x - 1)S_x^2 + (n_y - 1)S_y^2}{n_x + n_y - 2}}$
p-value: p-value $= P\left(T > \frac{\bar{x} - \bar{y}}{S_D\sqrt{1/n_x + 1/n_y}}\right) < \alpha$, rej, (D.F: $n_x + n_y - 2$)
where pooled $S_D = \sqrt{\frac{(n_x - 1)S_x^2 + (n_y - 1)S_y^2}{n_x + n_y - 2}}$
3. right-sided: $H_0: \mu_x = \mu_y$ vs. $H_1: \mu_x < \mu_y \longrightarrow H_0: \mu_y - \mu_x = 0$ vs. $H_1: \mu_y - \mu_x > 0$: change $\bar{x} - \bar{y}$ to $\bar{y} - \bar{x}$ on each
4. special case: each pair of X_i, Y_i dependent. paired t-test
①: prob desc. $X_i \sim N(\mu_x, \sigma_x^2)$, $Y_i \sim N(\mu_y, \sigma_y^2)$. X_i, Y_i (each pair) dependent; $H_0: \mu_x = \mu_y$ vs. $H_1: \mu_x \ne \mu_y$
②: analysis: can't use critical value / p-value: assume idpd; can: $X - Y \sim N(\mu_x - \mu_y, \sigma_x^2 + \sigma_y^2 - 2\sigma_{xy})$
→. change the prob to: let $D_i = Y_i - X_i$, $D \sim N(\mu_d, \sigma_D^2)$: $H_0: \mu_d = 0$ vs. $H_1: \mu_d \ne 0$; σ_D^2 unknown
~~where~~ $\cancel{\mu_d = \mathbb{E}[D_i] = \mathbb{E}[Y_i - X_i] = \mu_x - \mu_y}$,
③: sol. <i> let $D_i = Y_i - X_i$, find each pair's D_i <ii> write: "$D \sim N(\mu_d, \sigma_D^2)$ $H_0: \mu_d = 0$, $H_1: \mu_d \ne 0$)
<iii> Ⅳ - 2 sided

Ⅵ: other notes

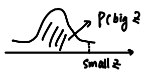
1. ☆: ①. find spl, asmp: assume ... is X, ... is Y; $H_0 \cdots$, $H_1 \cdots$; ②. find suitable type & situ ③. cal & compare
2. asking for p in bernoulli same with asking for μ; use the same way: when happen is 1, when not happen is 0
3. ☆: CCLD on Bernoulli variance
①: interval testing, σ^2 unknown Bernoulli: $S = \sqrt{\bar{x}(1-\bar{x})}$
②: hypo testing, only one population; σ^2 "unknown" (not explicitly said), Bernoulli: actual σ^2 known.
σ can be known from hypo: $\sigma = \sqrt{\mu_0(1-\mu_0)}$
③. hypo testing, 2 ppl σ^2 unknown Bernoulli
($S_x^2 = S_y^2$, can't be inferred from H_0)
$\Rightarrow S_p = \sqrt{\hat{p}(1-\hat{p})}\sqrt{\frac{1}{n_x} + \frac{1}{n_y}}$, where $\hat{p} = \frac{\sum_{i=1}^{n_x} x_i + \sum_{i=1}^{n_y} y_i}{n_x + n_y}$ # of 1 in x / y

CDF table of standard normal distribution $P(Z \le z)$

Small z

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

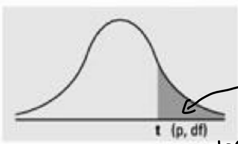
big $z = \frac{x-\mu}{\sigma}$



pbb value

Numbers in each row of the table are values on a t-distribution with (df) degrees of freedom for selected right-tail (greater-than) probabilities (p).

$T = \frac{X-\mu}{s}$



P(T≥t)

t number

df/p	0.40	0.25	0.10	0.05	0.025	0.01	0.005	0.0005
1	0.324920	1.000000	3.077684	6.313752	12.70620	31.82052	63.65674	636.6192
2	0.288675	0.816497	1.885618	2.919986	4.30265	6.96456	9.92484	31.5991
3	0.276671	0.764892	1.637744	2.353363	3.18245	4.54070	5.84091	12.9240
4	0.270722	0.740697	1.533206	2.131847	2.77645	3.74695	4.60409	8.6103
5	0.267181	0.726687	1.475884	2.015048	2.57058	3.36493	4.03214	6.8688
6	0.264835	0.717558	1.439756	1.943180	2.44691	3.14267	3.70743	5.9588
7	0.263167	0.711142	1.414924	1.894579	2.36462	2.99795	3.49948	5.4079
8	0.261921	0.706387	1.396815	1.859548	2.30600	2.89646	3.35539	5.0413
9	0.260955	0.702722	1.383029	1.833113	2.26216	2.82144	3.24984	4.7809
10	0.260185	0.699812	1.372184	1.812461	2.22814	2.76377	3.16927	4.5869
11	0.259556	0.697445	1.363430	1.795885	2.20099	2.71808	3.10581	4.4370
12	0.259033	0.695483	1.356217	1.782288	2.17881	2.68100	3.05454	43178
13	0.258591	0.693829	1.350171	1.770933	2.16037	2.65031	3.01228	4.2208
14	0.258213	0.692417	1.345030	1.761310	2.14479	2.62449	2.97684	4.1405
15	0.257885	0.691197	1.340606	1.753050	2.13145	2.60248	2.94671	4.0728
16	0.257599	0.690132	1.336757	1.745884	2.11991	2.58349	2.92078	4.0150
17	0.257347	0.689195	1.333379	1.739607	2.10982	2.56693	2.89823	3.9651
18	0.257123	0.688364	1.330391	1.734064	2.10092	2.55238	2.87844	3.9216
19	0.256923	0.687621	1.327728	1.729133	2.09302	2.53948	2.86093	3.8834
20	0.256743	0.686954	1.325341	1.724718	2.08596	2.52798	2.84534	3.8495
21	0.256580	0.686352	1.323188	1.720743	2.07961	2.51765	2.83136	3.8193
22	0.256432	0.685805	1.321237	1.717144	2.07387	2.50832	2.81876	3.7921
23	0.256297	0.685306	1.319460	1.713872	2.06866	2.49987	2.80734	3.7676
24	0.256173	0.684850	1.317836	1.710882	2.06390	2.49216	2.79694	3.7454
25	0.256060	0.684430	1.316345	1.708141	2.05954	2.48511	2.78744	3.7251
26	0.255955	0.684043	1.314972	1.705618	2.05553	2.47863	2.77871	3.7066
27	0.255858	0.683685	1.313703	1.703288	2.05183	2.47266	2.77068	3.6896
28	0.255768	0.683353	1.312527	1.701131	2.04841	2.46714	2.76326	3.6739
29	0.255684	0.683044	1.311434	1.699127	2.04523	2.46202	2.75639	3.6594
30	0.255605	0.682756	1.310415	1.697261	2.04227	2.45726	2.75000	3.6460
z	0.253347	0.674490	1.281552	1.644854	1.95996	2.32635	2.57583	3.2905
CI	———	———	80%	90%	95%	98%	99%	99.9%

§1. pbb.

△copy, cal, think neatly

§1.1 pbb by counting.

I. basic def.
1. sample space (Ω), event, experiment
2. cartinality: number of possible outcomes $|\Omega|$

II. pbb by counting
1. discrete pbb law: add up to 1
2. discrete uniform pbb law: $P(A)=\frac{|A|}{|\Omega|}$ △all with same pbb

III. counting techniques.
1. product rules: if event "linear", num of outcomes = product of event outcomes
① think: "tree"; "?×?×?"
② use. <1>. find $|A|$; <2>. permutation (※ in test write "!" not $P(\dots)$")
2. combination
①. binomial coefficient (combination) $\binom{n}{k}=\frac{n!}{k!(n-k)!}$
② multinomial coefficient. $\binom{n}{k_1,k_2,\dots k_t}=\frac{n!}{k_1!k_2!\cdots k_t!}$, where $k_1+\dots+k_t=n$
→. a whole group to several sub-groups

§1.2 probability axioms. (relation between events, cal pbb from pbb)
1. def. ① complement, $\cap$, $\cup$, ②. disjoint / mutually exclusive: $A\cap B=\emptyset$
2. pbb axioms ①. $P(A)\geq 0$ ② $P(\Omega)=1$ ③. if A_1, A_2 disjoint, $P(A_1\cup A_2\cup\dots)=P(A_1)+P(A_2)+\dots$
→. from these axioms can get everything from Venn graph
3. commonly used rules. (can get from axioms / venn graph)
① (sum rule) pbb of finite spaces. pbb = sum of all disjoint outcomes
②. $P(E^c)=1-P(E)$ ③. $P(E\cup F)=P(E)+P(F)-P(E\cap F)$
④. $A\cup B\cup C=A+B+C-(A\cap B+B\cap C+A\cap C)+A\cap B\cap C$

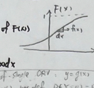
§1.3 conditional pbb. (for conditional events also multi-events pbb)
I. $P(A|F)=\frac{P(A\cap F)}{P(F)}=\frac{|A\cap F|}{|F|}$ (counting)
II. pbb axiom in conditional: ①: $P(A|F)\geq 0$ ② $P(F|F)=1$ ③. disjoint. $P(A\cup B|F)=P(A|F)+P(B|F)$
→. other rules from venn graph…
III: multiplication rule (pbb): $P(E_1\cap E_2)=P(E_1)\cdot P(E_2|E_1)$
→ general: $P(E_1\cap E_2\cap E_3)=P(E_1)\cdot P(E_2|E_1)\cdot P(E_3|E_2\cap E_1)\cdots$
IV: total pbb:
$P(E)=P(E|F)\cdot P(F)+P(E|F^c)\cdot P(F^c)$; → generalized: $\sum_i P(E|x_i)\cdot P(x_i)$
V: Bayes rule: $P(A|B)=\frac{P(B|A)\cdot P(A)}{P(B)}=\frac{P(B|A)\cdot P(A)}{P(B|A)\cdot P(A)+P(B|A^c)\cdot P(A^c)}$ → generalized

§1.4. idpd
I. basic:
1. def: $P(A\cap B)=P(A)\cdot P(B)\Leftrightarrow P(A)=P(A|B)\Leftrightarrow P(B)=P(B|A)$
2. ppt. ①: if E, F idpd; E^c, F idpd
II. idpd of multiple events.
1. def: all subsets of the events; pbb of intersection = mul of pbb
→ 3 events, idpd ⇔ AB, BC, AC idpd and $P(A\cap B\cap C)=P(A)\cdot P(B)\cdot P(C)$
2. ppt: ①: generalized: all complements are independent.
III. idpd of conditional events.
1. def: $P(A\cap B|F)=P(A|F)\cdot P(B|F)\Leftrightarrow P(A|B\cap F)=P(A|F)\Leftrightarrow\cdots$
2. ppt ① A|F, B|F idpd ⇒ $A^c|F$, B|F idpd and ~~A|F^c, B|F^c idpd~~ not
②: conditional idpd ⇎ idpd

§1.5. ※ in pbb
1. find pbb of normal event:
①, counting and $P(A)=\frac{|A|}{|\Omega|}$ ② sum rule / mul rule (pbb) ③. total pbb thm.
2. find pbb related to multiple events.
① $\cup$: sum rule ② $\cap$: mul rule. $P(A)\cdot P(B|A)$: don't know idpd; $P(A)\cdot P(B)$ idpd
③ venn graph
3. pbb for conditional events.
①: conditional formula: by pbb / counting ②: Baye's rule
③ by simply understand: E|F then F is Ω…
④. my formula: $(B_1|B_2)|B_3=B_1|B_2\cap B_3$

§4: ※ in RV. draw PDF and …
1. find $f_X(x)$
sv: { ①. uniform: $f(x)$ same in interval, $\int_{-\infty}^{\infty}f(x)dx=1$: cal by integation / geometry
②. derivative from $F(X)$ / $P(X)$
③. if fits common model }
mv { ①: get marginal: $f_X(x)=\int_{-\infty}^{\infty}f_{X,Y}(x,y)dy$
②: Total pbb: $f_X(x)=\int_{-\infty}^{\infty}f_{X|Y}(x|y)\cdot f_Y(y)dy$ }
single der: $f_X(x)=\frac{\partial}{\partial y}F_{X,Y}$
2. find joint $f_{X,Y}(x,y)$
①. uniform: $\iint f\,dxdy=1$: solve c
②: double der: $\frac{\partial}{\partial x}\frac{\partial}{\partial y}F_{X,Y}(x,y)$
3. find $F(x)$
sv { ①: from def: $F(x)=P(X\leq x)$
② itg of $f(x)$
③ from model }
4. joint $F_{X,Y}(x,y)$
①. from def
② itg of $f_{X,Y}(x,y)$
5. find pbb $P(a\leq X\leq b)$
①: from $F(x)$ / $F_{X,Y}(x,y)$
②. itg from $f(x)$, $f_{X,Y}(x,y)$
③. use pbb rules from §1: $a\leq X\leq b$ as an event
6. find conditional PDF:
①. def. formula
②: continuous Bayes
③: $f_{X|X\geq a}=\frac{\frac{\partial}{\partial x}P(a\leq X\leq x)}{P(X\geq a)}$
7. find $E(X)$
①: def: $E(X)=\int_{-\infty}^{\infty}xf(x)dx$ ⑪ by actual meaning: eg geometric. $E(X|X>1)\cdots$)
②: common model
③. from fn of MV: $E(g(x,y))=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g(x,y)f_{X,Y}(x,y)dxdy$
④. lineanility: $E(\alpha X+\beta Y)=\alpha E(X)+\beta E(Y)$
⑤. uncrl: $E(XY)=E(X)E(Y)$
⑥. idpd: $E(h(X)g(Y))=E(h(X))E(g(Y))$
⑧. total Expectation, continuous or discrete. $E(X|X=1)P(X=1)+E(X|X>1)\cdot P(X>1)$
⑨. $E(X^2)=Var(X)+E(X)^2$
⑩. indicator: $E(X)=E(X_1)+\dots+E(X_n)$
8. find $Var(X)$
①. def
②: $E(X^2)-E(X)^2$
③, get from common model
④: uncorrelated: $Var(X+Y)=Var(X)+Var(Y)$
⑤. rules of var: $Var(a+X)=Var(X)$, $Var(bX)=b^2Var(X)$
⑥. indicator: <1> get $E(X)$ <2>. $E(X^2)=E(\sum_{i=1}^{n}x_i^2+\sum_{\substack{i\neq j\\1\leq i,j\leq n}}x_ix_j)$ ← (n^2-n) in total
9. find $Cov(X,Y)$
①. def
②: $E(XY)-E(X)E(Y)$
10. Normal approximation: ①: sum / mean: central limit; ②: Binomial / Poisson
11. with both f and P: write f as P (use pbb bayes (eg). the eliminate dx from P to get f
→. if still don't know, see PMF as PDF

§3. continuous RV - def - diff - func △: $f(x)$ / $F(X)=0, 1$ otherwise
§3.1 defs (new)
I. Non-discrete pbb model
II. CDF: $F(x)=P(X\leq x)$
1. ppt: ①: $F(x)$ strictly increases ②: $\lim_{x\to\infty}F(x)=1$; $\lim_{x\to-\infty}F(x)=0$
III. PDF: probability density fn $f(x)$
1. def: $f(x)=\lim_{\delta\to 0}\frac{F(x+\delta)-F(x)}{\delta}=\frac{dF(x)}{dx}$ (tangent of $F(x)$)

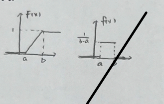


2. ppt: ①: PDF ≥ 0, and can be any num, eg: >1
②: integration of PDF is pbb: sum of $f(x)dx$
III: pbb of CRV.
1. $P(X=x)=0$
2. $P(E)=\int_E f(x)dx$ (E is interval in R)
3. $P(X\leq t)=\int_{-\infty}^{t}f(x)dx$
4. $\int_{-\infty}^{\infty}f(x)dx=1$

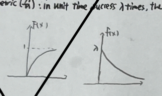
Func of single CRV: $y=g(x)$
①: $F_Y(y)$: by def. $P(Y\leq y)=P(X\leq\cdots)$
②: $f_Y(y)=\frac{d}{dy}F_Y(y)$

IV: expected value. $E(X)=\int_{-\infty}^{\infty}x\,f(x)dx$ → Expectation of function $g(x)$: $E(g(x))=\int_{-\infty}^{\infty}g(x)f(x)dx$
V: variance: $Var(X)=E(X^2)-E(X)^2=\int_{-\infty}^{\infty}x^2f(x)dx-(\int_{-\infty}^{\infty}xf(x)dx)^2$

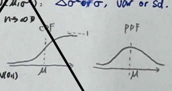
§3.2. diff CRV
I. uniform random var: $T\sim$ Uniform(a,b)
1. CDF: $F(t)=P(X\leq t)=\begin{cases}0 & (t<a)\\ \frac{1}{b-a}t-\frac{a}{b-a} & (a\leq t\leq b)\\ 1 & (t>b)\end{cases}$



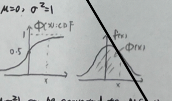
2. PDF: $f(t)=\begin{cases}0 & (t<a)\\ \frac{1}{b-a} & (a\leq t\leq b)\\ 0 & (t>b)\end{cases}$
3. $E(X)=\frac{a+b}{2}$ 4. $Var(X)=\frac{(b-a)^2}{12}$
II. exponential RV: Exponential (λ)
1. def. Exponential(λ) = $\lim_{n\to\infty}\frac{1}{n}$Geometric($\frac{\lambda}{n}$): in unit time success λ times, the time of 1st success
2. PDF / def.
$f(x)=\begin{cases}0 & (x<0)\\ \lambda e^{-\lambda x} & (x\geq 0)\end{cases}$



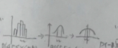
3. CDF: $F(x)=1-e^{-\lambda x}$
4. $E(X)=\frac{1}{\lambda}$; 5. $Var(X)=\frac{1}{\lambda^2}$
6. ppt: memoryless, past has no effect on present: $P(X\geq s+t|X\geq t)=P(X\geq s)$
III. Normal (Gaussian) distribution. $N(\mu,\sigma^2)$: △σ^2 or σ, var or sd.
1. think (Binomial (actually all) for $n\to\infty$)
2. PDF / def:
$f(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$



3. CDF: integration: look up table of N(0,1)
4. $E(X)=\mu$; 5: $Var(X)=\sigma^2$
IV: Standard normal distribution. N(0,1)
1. def: zero mean and unit variance: $\mu=0$, $\sigma^2=1$
2. PDF: $f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$
3. CDF: $\Phi(x)$: look up table
→ ppt: $\Phi(x)=1-\Phi(-x)$ (draw)

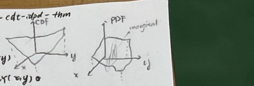


4. ppt ①: relation with $N(\mu,\sigma^2)$: all $N(\mu,\sigma^2)$ can be converted to N(0,1)
suppose $X\sim N(\mu,\sigma^2)$; $Z=\frac{X-\mu}{\sigma}$ is N(0,1) ← for all RV: can normalize this way: $\mu_Z=0$ and $\sigma_Z^2=1$
5. N(0,1) and Binomial
①: foundamental. DeMoivre Laplace Limit Thm.



$S_n\sim$ Binomial(n, p) ⇒ $\mu=np$, $\sigma^2=np(1-q)$; normalize $S_n\to Z_n=\frac{S_n-np}{\sqrt{np(1-p)}}$:
$\lim_{n\to\infty}(a\leq Z_n=\frac{S_n-np}{\sqrt{np(1-p)}}\leq b)=\Phi(b)-\Phi(a)$ (CDF of Binomial converge to CDF of $N(\mu,\sigma^2)$)
②: ※: normal approximation of Binomial / Poisson: see CDF of Binomial as Normal
<1>. $Z=\frac{X-np}{\sqrt{np(1-p)}}\sim N(0,1)$
<2>. $P(a\leq X\leq b)=P(a-0.5\leq X\leq b+0.5)=P(a'\leq Z\leq b')=\Phi(a')-\Phi(b')$
eg: $P(45\leq X\leq 55)=P(44.5\leq X\leq 55.5)$: each num a in Binomial is (a-0.5, a+0.5)

§3A mul-var CRV - def - (diff) - func - cdt - idpd - thm



§3A-1 new defs:
I: joint and marginal
1. { joint CDF: $F_{X,Y}(x,y)=P(X\leq x, Y\leq y)$
joint PDF: $f_{X,Y}(x,y)=\frac{\partial}{\partial x}\frac{\partial}{\partial y}F_{X,Y}(x,y)$ }
2. { marginal CDF: $F_X(x)=P(X\leq x)=\lim_{y\to\infty}F_{X,Y}(x,y)$
marginal PDF: $f_X(x)=\int_{y=-\infty}^{\infty}f_{X,Y}(x,y)dy$ (= $\int_{-\infty}^{\infty}f_X(x)\cdot f(Y=y|X=x)dy$) }
II. pbb of mulvar CRV: volumn under PDF curve
let A be event $\in R^2$, pbb: $\iint_A f(x,y)dxdy$
§3A-2 diff kinds of mul-var CRV
I. "higher dimension uniform" over S
1. $f_{X,Y}(x,y)=\begin{cases}c & \text{if }(x,y)\in S\\ 0 & \text{otherwise}\end{cases}$
→ ppt: $\iint_S f_{X,Y}(x,y)=1$: to find c
2. $F_{X,Y}(x,y)$: integration
3. pbb: for $A\in R^2$: $P(A)=\iint_A f_{X,Y}(x,y)dxdy$
4. situation: (no need) X, Y idpd uniform RV, then in [X]×[Y] region…
①: same pbb for every points to a region → just use geometric formula to get $F(x)$ derivative to get $f(x)$ ↳ just $f_{X,Y}=f_X\cdot f_Y$
§3A-3 fn of mul-var CRV: $Z=g(X,Y)$
(1) I: CDF of Z: get by meaning, $F_Z(z)=P(Z\leq z)$ and cal / use integration / volumn
(2) II. PDF of Z: ①: $f_Z(z)=\frac{\partial}{\partial z}F(z)$ ② if uncrl and $g(x,y)=xy$, f
III. $E(Z)=E(g(X,Y))=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g(x,y)f_{X,Y}(x,y)dxdy$
§3A-4 conditional MVCRV
I. conditional PDF
1. def: $f_{X|Y}(x|y)=\frac{f_{X,Y}(x,y)}{f_X(x)}$
→. for the other kind: $\frac{f_X(x)}{P(E)}$ (eg $f_{X|X\geq a}(x)=\frac{\frac{d}{dx}P(a\leq X\leq x)}{P(X\geq a)}=\frac{f_X(x)}{P(X\geq a)}$)
II. cdt CDF: $F_{X|E}(x)=P(X\leq x|E)$: by pbb / itg
III. cdt expectation
1. def: $E(X|B)=\int_B x f_{X|B}(x)dx$ → eg: $\begin{cases}E(X|Y=y)=\int_{-\infty}^{\infty}x\cdot f_{X|Y}(x|y)dx\\ E(X|X\geq a)=\int_a^{\infty}xf_{X|X\geq a}(x)dx\end{cases}$
IV: formulas related:
1. total: ①. total PDF: $f_X(x)=\int_{y=-\infty}^{\infty}f_{X|Y}(x|y)f_Y(y)dy$
②: total expectation: $E(X)=\int_{-\infty}^{\infty}E(X|Y=y)\cdot f_Y(y)dy$
③: total pbb: $P_X(x)=\int_{-\infty}^{\infty}P(X=x|Y=y)\cdot f_Y(y)dy$
2. continuous Bayes rule: $f_{X|Y}(x|y)=\frac{f_{Y|X}(y|x)\cdot f_X(x)}{f_Y(y)}$
§3A-5. idpd in Mul-var CRV
I. idpd 1. def: X, Y idpd ⇔
mul { ①: $F_{X,Y}(x,y)=F_X(x)\cdot F_Y(y)$ or
②: $f_{X,Y}(x,y)=f_X(x)\cdot f_Y(y)$ }
2. conditional { ①: $F_{X|Y}(x|y)=F_X(x)$ or
②: $f_{X|Y}(x|y)=f_X(x)$ }
(3): any g, h $E(g(X)h(Y))=E(g(X))E(h(Y))$
II. uncorrelated: 1. def: X, Y ucrl ⇔ $E(XY)=E(X)\cdot E(Y)$; ucrl ⇐ idpd
§3A-6 Central limit thm: n large: discrete → $N(\mu,\sigma^2)$ ($\frac{\bar{X}_n-E(\bar{X}_n)}{\sqrt{Var(\bar{X}_n)}}$ is N(0,1))
1. def: IID RV $x_1, x_2, \cdots x_n$ with ① μ, σ^2, define ② $Z_n=\frac{\frac{1}{n}(x_1+x_2+\dots+x_n)-\mu}{\frac{\sigma}{\sqrt{n}}}$
the CDF of Z_n coverge to CDF of N(0,1): $\Phi(X)$:
$\lim_{n\to\infty}P(Z\leq z)=\Phi(z)$
2. ※ ①: let $x_i\cdots$, find μ, σ^2 ②: $S_n=x_1+\dots+x_n$ ③ $Z_n=\frac{\frac{1}{n}\overbrace{(x_1+\dots+x_n)}^{S_n}-\mu}{\frac{\sigma}{\sqrt{n}}}\sim N(0,1)$
④ $P(S\in\cdots)=P(Z\in\cdots)=\cdots$

§2. Random Var (Discrete) def - diff - func - thm
§2.1 basic def:
I. discrete RV
II. PMF: probability mass fn.
1. def: all the $P(X=\dots)=?$ ②: describle: table / chart // test: $P(X=\dots)$ for all: ok
3. ppt: ①: $x\in R$. $P(X\geq 0)$ ②. add up to 1
III. CDF: cumulative distribution fn: $F(x)=P(X\leq x)$
IV: expectation: $E(X)=\sum_x xp(x)$: average value of RV when many experiments
→. gain in games → $E(const)=const$
V: variance. how close x and E(x) are
$Var(X)=E[(X-\mu)^2]=\sum_{\text{all }x}(x-E(x))^2p(x)$ ($\mu=E(X)$)
→. or: $Var(X)=E(X^2)-E^2(X)$
→ ppt: $Var(X+a)=Var(X)$; $Var(bX)=b^2Var(X)$
VI. Standard deviation: $\sigma=\sqrt{Var(X)}$
§2.2 diff RV
0. Bernouli
1. def: Bernoulli(p) = Binomial(1, p): only 0, 1
2. PMF: $P(X=0)=1-p$; $P(X=1)=p$ 3. $E(X)=p$
I. Binomial:
1. def: X is num of success where: try n times, each time pbb of success: p
$X\sim$ Binomial(n, p)
2. PMF. $P(X=k)=\binom{n}{k}\cdot p^k(1-p)^{n-k}$ 3. CDF: just add up
4. $E(X)=np$ (n× Bernoulli(p)) 5. $Var(X)=np(1-p)$
II. geometric
1. def: X is num of trial until first success. each time success pbb = p
$X\sim$ Geometric(p)
2. PMF. $P(X=k)=(1-p)^{k-1}p$ 3. CDF. $P(X\leq k)=1-(1-p)^k$ ($\sum_{i=1}^{\infty}ir^{i-1}=\frac{1}{(1-r)^2}$ (0<r<1))
4. $E(X)=\frac{1}{p}$ 5. $Var(X)=\frac{1-p}{p^2}$
III: poisson:
1. def. Poisson(λ): Binomial($n,\frac{\lambda}{n}$) when $n\to\infty$: distribution when expected times is λ (and events can be separated to extremely small): $X\sim$ poisson(λ)
2. PMF: $P(X=k)=e^{-\lambda}\frac{\lambda^k}{k!}$ 3. CDF: just add up ($\lim_{n\to\infty}(1+\frac{x}{n})^n=e^x$)
4. $E(X)=\lambda$ 5. $Var(X)=\lambda$
§2.3 fn of RV: if 2 RV are related (mapped by fn $Y=f(X)$), how to get Y from X?
I. PMF of $Y=f(X)$:
$P_Y(y)=\sum_{x:f(x)=y}P_X(x)$ (sum of all pbb of all situations of y)
II. expectation of ~
$E(f(X))=\sum_x f(x)\cdot p_X(x)$
→ $E(f(X))\neq f(E(X))$
§2.4. thm about inequality:
I. markov's inequality: suppose RV $X\geq 0$ (non-negative only), for every $t>0$
$P(X\geq t)\leq\frac{E(X)}{t}$ ($P(X\geq t)\cdot t\leq P(X\geq t)\cdot E(X|X\geq t)\leq E(X)$)
II. chebyshev's inequality. suppose X, μ, σ^2, for every $t>0$
$P(|X-\mu|\geq t)\leq\frac{\sigma^2}{t^2}$ cor: $P(|X-\mu|\geq S\sigma)\leq\frac{1}{S^2}$ (let $Y=(X-\mu)^2\geq 0$, use markov's…)
§2.5. Law of large numbers: further explanation about expectation. mean of many x_i is E(X)
I. empirical mean $\bar{X}_n$ of idpd and identicaly distributed (IID) RV: $x_1, x_2, \cdots x_n$, with μ, σ^2
2. $E(\bar{X}_n)=\mu$ ($E(\frac{x_1}{n})+E(\frac{x_2}{n})+\dots+E(\frac{x_n}{n})=n\cdot\frac{\mu}{n}$) $\bar{X}_n=\frac{x_1+x_2+\dots+x_n}{n}$
3. $Var(\bar{X}_n)=\frac{\sigma^2}{n}$ ($Var(\frac{x_1}{n}+\frac{x_2}{n}+\dots\frac{x_n}{n})=Var(\frac{x_1}{n})+Var(\frac{x_2}{n})+\dots+Var(\frac{x_n}{n})=n\cdot\frac{\sigma^2}{n^2}=\frac{\sigma^2}{n}$)
II. LLN:
1. def. $\lim_{n\to\infty}\bar{X}_n=\lim_{n\to\infty}\frac{x_1+x_2+\dots+x_n}{n}=\mu$ (proof: chebyshev: $P(|\bar{X}_n-\mu|\geq\epsilon)\leq\frac{\sigma^2}{n\epsilon^2}$ sandwich
2. W and S LLN:
IID $x_1, x_2\dots x_n$: μ and finite variance σ^2, for every $\epsilon>0$
{ WLLN: $\lim_{n\to\infty}P(|\bar{X}_n-\mu|\geq\epsilon)=0$
SLLN. $P(\lim_{n\to\infty}\bar{X}_n=\mu)=1$ } WLLN ⇐ SLLN

§2A. multivariable discrete RV - def - func - cdt - idpd
§2A.1 basic (newly defined) def.
I. def
1. joint PMF. $P_{X,Y}(x,y)=P(X=x, Y=y)$ joint ⇒ marginal
2. marginal PMF: $P_X(x)$, $P_Y(y)$: $P_X(x)=\sum_{y\in Y}P_{X,Y}(x,y)$
§2A.2 fn of multivariable (discrete) RV
→ when an event determined by multiple events ($Z=f(X,Y)$): get features of Z from X, Y
I. PMF: $P_Z(z)=\sum_{x,y\in f(x,y)=z}P_{X,Y}(x,y)$
II. expectation:
1. formula: $E(f(x,y))=\sum_{\text{all }x,y}f(x,y)\cdot P_{X,Y}(x,y)$
2. linearility of E: $E(\alpha X+\beta Y)=\alpha E(X)+\beta E(Y)$
§2A-3 conditional RV. when multi-var: can have conditional anything
I. conditional PMF: $P_{X|Y}(x|y)=P(X=x|Y=y)=\frac{P_{X,Y}(x,y)}{P_Y(y)}$
II. conditional CDF. $F_{X|E}(x)=P(X\leq x|E)$ eg: $F_{X|X\geq 0.5}(x)=P(X\leq x|X>0.5)$
III. conditional expectation.
1. def. $E(X|B)=\sum_{\substack{\text{all }x\text{ happens}\\ \text{in condition }B}}xP(X|B)$
2. common examples.
①: $E(X|Y=y)=\sum_{\text{all }x}xP(X=x|Y=y)=\sum_{\text{all }x}xP_{X|Y}(x|y)$
②. $E(X|X>1)=\sum_{\text{all }x>1}xP(X=x|X>1)$
3. formulea: total expectation
$E(X)=\sum_y P(Y=y)\cdot E(X|Y=y)$: several parts $\sum E(\text{part})\cdot P(\text{part})$
IV. formulas | total pbb, $P(X\in S)=\sum_y P(X\in S|Y=y)\cdot P(Y=y)$

	E	Var
$E(f(x))\neq f(E(x))$	linear	aX, bX
uncrl	$E(XY)=E(X)E(Y)$	$Var(X+Y)=Var(X)+Var(Y)$
idpd	$E(f(x)g(y))=E(f(x))E(g(y))$	

§2A-4 idpd in RV
I. idpd
1. def: X, Y idpd ⇔
①: every pair of x, y: $P(X=x, Y=y)=P(X=x)\cdot P(Y=y)$
②: for every outcome y, $P_{X|Y}(\cdot|y)=P_X(\cdot)$ / $P_{X|Y}(x|y)=P_X(x)$: for all x, y
③ every f, g. $E(f(x)g(y))=E(f(x))E(g(y))$
II. variance and uncorrelated.
1: covariance: how related are 2 RVs
① def. $Cov(X,Y)=E[(X-\mu_X)(Y-\mu_Y)]=E(XY)-E(X)E(Y)$
2. uncorrelated
①: def X, Y uncrl ⇔ $Cov(X,Y)=0$ ⇔ $E(XY)=E(X)E(Y)$
②: ppt <1> X, Y idpd ⇒ X, Y uncrl
<2>. X, Y uncrl ⇒ $Var(X+Y)=Var(X)+Var(Y)$
→ general: if $x_1, x_2\dots x_n$ pairwise idpd (every pair idpd):
$Var(x_1+x_2+\dots+x_n)=Var(x_1)+Var(x_2)+\dots+Var(x_n)$

CCLD: all shortened expressions.
1. single var. $P_X(x)=P(X=x)$; $f_X(x)=f(X=x)$ (also $f(x)$); $F_X(x)=P(X\leq x)$ (also $F(x)$)
2. mul var: ①. $P_{X,Y}(x,y)=P(X=x, Y=y)$; ②. $f_{X,Y}(x,y)=f(X=x, Y=y)$ ③ $F_{X,Y}(x,y)=P(X\leq x, Y\leq y)$
3. conditional,
①: $P_{X|Y}(x|y)=P(X=x|Y=y)$; $P_{X|E}(x)=P(X=x|E)$ eg: $P_{X|Y>3}(x)=P(X=x|Y=3)$
②: $f_{X|Y}(x|y)=f(X=x|Y=y)$; $f_{X|E}(x)=f(X=x|E)$
③. (?) $F_{X|Y}(x|y)=P(X\leq x|Y=y)$; $F_{X|E}(x)=P(X\leq x|E)$
④. conditional expectation: see up.

CCLD: common Integration:
1. $\int f(x)g(x)dx=F(x)g(x)-\int F(x)g'(x)dx$ 2. $\int e^{kx}dx=\frac{1}{k}e^{kx}$, $\int a^{kx}dx=\frac{1}{k}\frac{1}{\ln a}a^{kx}$ 3. $\int x^ndx=\frac{1}{n+1}x^{n+1}$ 4. $\int\frac{1}{x}dx=\ln|x|$
2. $(\frac{1}{x})'=-\frac{1}{x^2}$, $(\frac{1}{\sqrt{x}})'=\frac{1}{2\sqrt{x}}$, $(\log_a x)'=\frac{1}{x\ln a}$, $(a^x)'=\ln a\cdot a^x$, $(a^u)'=a^u(\ln a)u'$
$[f(x)g(x)]'=f'(x)g(x)+f(x)g'(x)$. $[\frac{f(x)}{g(x)}]'=\frac{f'(x)g(x)-f(x)g'(x)}{(g(x))^2}$
→. $P(Y>5|X=1)=\int_5^{\infty}f_{Y|X}(y|x)\cdot dy$
$f(X=1|Y>5)=\frac{\int_5^{\infty}f_{X,Y}(x,y)dy}{\int_5^{\infty}f_Y(y)dy}$

△: up to how many decimals
△: σ or σ^2, s or s^2!! , use formula: get is squared.

※: for hypo testing, eg. if all hypo. x impossible to be some num
just the range of x don't consider that num, $x\notin R$
also, when ÷0, can write ∞

$\int\sin(kx)dx=-\frac{1}{k}\cos(kx)$
$\int\cos(kx)dx=\frac{1}{k}\sin(kx)$
$\int\sec^2x\,dx=\tan x+C$, $\int\sec x\tan x\,dx=\sec x+C$
$\int\csc^2x\,dx=-\cot x+C$, $\int\csc x\cot x\,dx=-\csc x+C$
$\int\tan x\,dx=\ln|\sec x|+C$, $\int\cot x\,dx=-\ln|\csc x|+C$
$\int\sec x\,dx=\ln|\sec x+\tan x|+C$, $\int\csc x\,dx=-\ln|\csc x+\cot x|+C$
$\int\frac{1}{\sqrt{a^2-x^2}}dx=\arcsin(\frac{x}{a})+C$
$\int\frac{1}{a^2+x^2}dx=\frac{1}{a}\arctan(\frac{x}{a})+C$