ETC3400 Principles of Econometrics

ECONOMICS · MONASH UNIVERSITY · LEVEL 3

Full guide + ask Sia about ETC3400

SIDE 1/2 DEFINITIONS → FORMULAS → DECISION RULES → COMPARISONS EXAM REVISION SHEET S2 2026 · final 60% · asksia.ai/cheatsheet/monash-etc3400

0 · Exam map

COURSE FACTS

A1 | A2 20% | A3 | Final Exam 60%

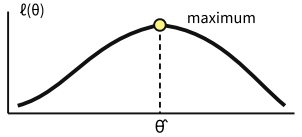
Assignment 1 10% · Assignment 2 20% · Assignment 3 10% · Final Exam 60%. The final is an individual examination on the e-assessment platform during the examination period.

Permitted materials: not confirmed for this offering in the course materials we hold — check the exam instructions on the LMS.

The final rewards rebuilding assignment-style reasoning: state the model, show the contribution to the likelihood, justify the theorem or restriction, carry algebra to the decision, then interpret. **Unsupported numbers are unfinished answers.**

1 · Likelihood → MLE

CH 1 · CORE METHOD



Likelihood profile: the score locates a candidate; curvature and boundaries certify it.

INDEPENDENT SAMPLE

$L(\theta;y)=\prod_i f(y_i;\theta)$ · $\ell(\theta)=\sum_i \log f(y_i;\theta)$

$U(\theta)=\partial\ell/\partial\theta$ · $H(\theta)=\partial^2\ell/\partial\theta^2$

1. Define sample space, parameter space and probability law.
2. Fix observed data; vary θ. Independence gives the product.
3. Take logs: strictly increasing, so the maximiser is preserved.
4. Solve $U(\theta)=0$ for an *interior candidate*.
5. Check negative curvature/global concavity, objective comparisons and every admissible boundary.

SEE	DO
factor depends only on observed y	may drop for maximisation
factor contains θ	must retain
open boundary approached	may be a supremum, not an MLE

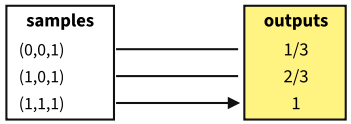
BERNOULLI REFERENCE DERIVATION

For twelve trials with eight successes, $L(\theta)\propto\theta^8(1-\theta)^4$ on $0\le\theta\le1$. Hence $\ell(\theta)=8\log\theta+4\log(1-\theta)$, $U(\theta)=8/\theta-4/(1-\theta)$ and $\hat\theta=8/12=2/3$. The second derivative $-8/\theta^2-4/(1-\theta)^2$ is negative for $0<\theta<1$; compare boundaries if the data make the interior formula hit an endpoint.

LINE	WHAT IT PROVES
log is increasing	same maximiser as positive likelihood
$U(\hat\theta)=0$	interior stationary candidate only

2 · Estimator logic

CH 2 · REPEAT SAMPLING



An estimator maps every possible random sample to one value; the realised output is the estimate.

OBJECT	CORRECT STATEMENT
Estimator T(Y)	random rule before observing data; has expectation, variance and sampling law
Estimate T(y)	fixed number after the realised dataset is inserted
sampling law	group identical outputs and add probabilities; multiplicity matters

COMPARISON CRITERIA

$\text{Bias}\theta(T)=E\theta[T]-g(\theta)$

$\text{MSE}\theta(T)=E\theta[(T-g(\theta))^2]=\text{Var}\theta(T)+\text{Bias}\theta(T)^2$

Use the same estimand and data-generating process. Unbiased does not mean low variance; a biased estimator can have smaller MSE when variance reduction exceeds squared bias. Unbiasedness is for *every* admissible θ, not one convenient value.

JENSEN DIRECTION

For convex g, $g(E[X]) \le E[g(X)]$; for concave g the inequality reverses. Identify curvature, write the direction, then substitute the known expectation.

"Nonlinear" alone does not identify upward or downward bias.

CLAIM	EVIDENCE REQUIRED
T is unbiased	show $E\theta[T]=g(\theta)$ for all admissible θ
T is more precise	smaller sampling variance under the same model
T has lower MSE	compare variance + squared bias pointwise in θ
T is consistent	show $T_n \to p$ target as $n\to\infty$
transform bias direction	name convex/concave g and write Jensen inequality

Procedure versus result: "T is unbiased" describes a repeated-sample rule. "The estimate is 0.6" describes one realised dataset. The estimator rule stays fixed as θ changes, but the model probabilities of possible datasets—and therefore the sampling distribution—change with θ.

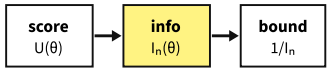
2b · Fast comparison

EXAM JUDGEMENT

QUESTION ASKS	CALCULATE / SAY
systematic displacement	bias
repeated-sample spread	variance
overall squared error	MSE
large-sample target	consistency, not finite-sample unbiasedness

3 · Information & bounds

CH 3 · PRECISION



Regular models: score variation and expected curvature measure local precision.

TWO REGULAR IDENTITIES

$I_n(\theta)=E\theta[U_n(\theta)^2]=-E\theta[\partial^2\ell_n(\theta)/\partial\theta^2]$

observed information $j_n(\theta)=-\partial^2\ell_n(\theta)/\partial\theta^2$

At the truth, $E\theta[U_n(\theta)]=0$ under regularity.

Independent scores add; zero expected-score cross terms vanish, so $I_n(\theta)=nI_1(\theta)$ for iid observations. The information identities may fail with parameter-dependent support or invalid differentiation under the integral.

SCALAR CRAMÉR–RAO

If $E\theta[T]=g(\theta)$, then $\text{Var}\theta(T) \ge [g'(\theta)]^2 / I_n(\theta)$

1. Name the target $g(\theta)$ and unbiased comparison class.
2. Check support, differentiation, finite information and identification conditions.
3. Derive the bound; then derive the candidate's variance.
4. Equality for all admissible θ ⇒ efficient *within that class/model*.

CLAIM	LIMIT
attains CR bound	minimum variance among relevant unbiased estimators
biased alternative	may still have lower MSE
vector covariance order	A−B positive semidefinite iff $c'(A-B)c\ge0$ for every c

REGULARITY GATE

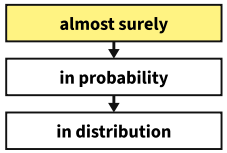
CHECK	WHY IT BITES
support	parameter-independent support underlies expected-score identities
differentiation	moving derivatives through the integral must be valid
score	zero mean and finite second moment
information	finite and nonsingular for local precision
target	identified and differentiable $g(\theta)$

For a fixed matrix A, a covariance comparison survives a linear target: if $\text{Var}(\tilde\beta)-\text{Var}(\hat\beta)$ is positive semidefinite, then $A[\text{Var}(\tilde\beta)-\text{Var}(\hat\beta)]A'$ is positive semidefinite because every quadratic form can be rewritten with $A'c$.

CENSORING TRAP → *Exact records use densities; right-censored records use survival probabilities. Treating the threshold as exact gives the wrong likelihood and information.*

4 · Convergence map

CH 4 · ASYMPTOTICS



Stronger modes imply weaker ones. Reverse arrows need extra conditions.

MODE	MEANING / USE
a.s.	sample paths converge outside a probability-zero set
p	probability of a material deviation vanishes; consistency uses this
d	CDFs approach a limiting law at continuity points; supports inference

DEFINITIONS TO RECOGNISE

$X_n \to p\ X$: for every $\varepsilon>0$, $P(|X_n-X|>\varepsilon)\to0$

$X_n \to d\ X$: $F_n(x)\to F(x)$ at continuity points of F

$X_n \to d\ c$ for constant c implies $X_n \to p\ c$. With a nonconstant limiting random variable, convergence in distribution alone generally does not imply convergence in probability.

MLE ASYMPTOTIC BRIDGE

$0=U_n(\hat\theta)=U_n(\theta_0)+U'_n(\tilde\theta)(\hat\theta-\theta_0)$

$\sqrt{n}(\hat\theta-\theta_0)=-[U'_n(\tilde\theta)/n]^{-1}[U_n(\theta_0)/\sqrt{n}]$

1. **Score zero** at the interior MLE.
2. **Taylor-expand** about the truth θ_0; label the intermediate $\tilde\theta$.
3. **CLT** supplies the score limit.
4. **LLN** supplies average Hessian convergence.
5. **Slutsky** combines them: typically $\sqrt{n}(\hat\theta-\theta_0) \to d\ N(0, I_1(\theta_0)^{-1})$.

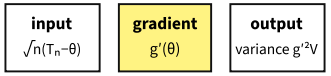
Asymptotic normality is not "estimators become normal"; it is a centred, scaled result requiring finite, nonsingular information and regularity.

STEP	QUANTITY CONTROLLED
consistency	$\tilde\theta$ lies near θ_0, so local curvature is relevant
CLT	random score $U_n(\theta_0)/\sqrt{n}$
LLN	average Hessian $U'_n(\tilde\theta)/n$
Slutsky	product of convergent score and inverse-curvature pieces

Do not reverse implications casually: almost-sure convergence controls paths and therefore yields probability convergence; probability convergence yields distributional convergence. Distributional convergence supplies a limiting law but not, in general, closeness of X_n and X on the same probability space.

5 · Delta methods

CH 4 · TRANSFORMS



First-order delta method: transform the centre and sandwich the variance by derivatives.

SCALAR

$\sqrt{n}(T_n-\theta)\to d\ N(0,V)$

$\Rightarrow \sqrt{n}[g(T_n)-g(\theta)]\to d\ N(0,[g'(\theta)]^2V)$

VECTOR

$\sqrt{n}(T_n-\theta)\to d\ N(0,V)$

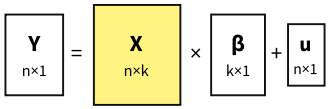
$\Rightarrow \sqrt{n}[g(T_n)-g(\theta)]\to d\ N(0,G\ V\ G')$, $G=\text{Jacobian at }\theta$

Write three separate lines: transformation, derivative/Jacobian, evaluation point. If $g'(\theta)=0$, the first-order approximation degenerates; use a higher-order expansion or another method, not "zero variance".

QUESTION WORDING	TOOL	OUTPUT
show $g(T_n)$ is consistent	continuous mapping	probability limit $g(\theta)$
approximate distribution / SE	delta method	centred, scaled limit and variance
$g'(\theta)=0$	higher-order expansion	do not use degenerate first-order answer

6 · Vector likelihood

CH 5 · DIMENSIONS



Write dimensions before multiplication; invalid algebra becomes visible immediately.

NORMAL LINEAR MODEL

$Y=X\beta+u$, $u\sim N(0,\sigma^2I_n)$, $\text{rank}(X)=k$

$\hat\beta=(X'X)^{-1}X'Y$ · $\text{Var}(\hat\beta)=\sigma^2(X'X)^{-1}$

VECTOR LIKELIHOOD OBJECTS

$U(\theta)=\partial\ell(\theta)/\partial\theta$ is $p\times1$

$H(\theta)=\partial^2\ell(\theta)/\partial\theta\partial\theta'$ and $I(\theta)=-E[H(\theta)]$ are $p\times p$

PROJECTION GEOMETRY

$P=X(X'X)^{-1}X'$ · $M=I-P$

$\hat Y=PY$ · $e=MY$ · $X'M=0$

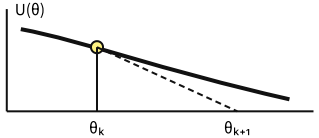
$P=P'=P^2$ · $M=M'=M^2$ · $\text{rank}(M)=n-k$

Full column rank makes $X'X$ invertible. Zero covariance between $\hat\beta$ and residuals becomes independence only because both are linear transformations of jointly normal Y. Zero covariance alone is not enough.

VARIANCE ESTIMATE	EXPECTATION
$\sigma^2_{ML}=e'e/n$	$((n-k)/n)\sigma^2$; bias $=-k\sigma^2/n$

7 · Numerical likelihood

CH 6 · AUDIT OUTPUT



Newton–Raphson: a tangent approximation to the score predicts its next root.

SCALAR / VECTOR UPDATE

$\theta_{k+1}=\theta_k-U(\theta_k)/U'(\theta_k)$

$\theta_{k+1}=\theta_k-H(\theta_k)^{-1}U(\theta_k)$

General routines minimise $-\ell(\theta)$; keep the sign convention consistent. BFGS builds a curvature approximation from objective and gradient information. Newton is local: poor starts, scaling, flat curvature or multiple basins can change the path.

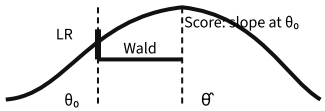
AUDIT	EVIDENCE TO RECORD
problem	data summary, parameter order, constraints, minimise/maximise sign
path	algorithm, starts, bounds, evaluations
ending	code, objective, final gradient, Hessian/curvature
cross-check	multiple starts; analytic, joint and concentrated fits

Positive scale: set $\lambda=\log\sigma$, optimise over real λ, return via $\sigma=\exp\lambda$. Compare bounded σ and log-scale runs only after transforming to a common parameterisation; estimates and maximised likelihood should match within numerical tolerance.

TRAP → *A successful convergence code reports how the routine stopped; it does not certify the model, derivatives or global optimum.*

8 · LR · Wald · Score

CH 7 · CHOOSE EVIDENCE



One likelihood, three views: height drop, scaled distance, and null slope.

TEST	FORMULA / EVALUATION
LR	$2[\ell(\hat\theta)-\ell(\tilde\theta)]\ge0$; needs both fits
Wald	$(R\hat\theta-r)'[R\hat V R']^{-1}(R\hat\theta-r)$; unrestricted fit
Score/LM	$U(\tilde\theta)'I(\tilde\theta)^{-1}U(\tilde\theta)$; constrained/null fit

Under regularity, each is asymptotically $\chi^2 q$, where q is the number/rank of independent restrictions, *not* the unrestricted parameter count. Known-variance normal mean: constant quadratic curvature gives exact equality. With estimated variance, the guide's profile comparison gives $LM \le LR \le Wald$.

Forward this to your tutorial group

ETC3400 EXAM REVISION SHEET

Monash Semester 2 final assessment period, 2–18 November 2026 (check your personal exam timetable). **Check your exam instructions on the LMS for what you may bring; use this sheet to revise and self-test.**

ETC3400 Principles of Econometrics

ECONOMICS · MONASH UNIVERSITY · LEVEL 3

Full guide + ask Sia about ETC3400

SIDE 2/2 FIVE WORKED DRILLS → JUDGEMENT CARDS → CLASSIC TRAPS → EXAM-MORNING PLAN SELF-TEST S2 2026 · show every answer path · asksia.ai/cheatsheet/monash-etc3400

9 · Drill 1: Poisson MLE

CH 1 WORKED

Given: nine independent Poisson counts, total 27. Derive the rate MLE and verify it.

joint L	log	score	verify
∏ pmf	sum	=0	ℓ″<0

Full-credit answer path: construct, simplify only constants, solve, then verify.

$$L(\lambda)=\exp(-9\lambda)\,\lambda^{27}\,/\,\textstyle\prod_i y_i!$$
$$\ell(\lambda)=-9\lambda+27\log\lambda+\text{constant}$$

$$U(\lambda)=-9+27/\lambda=0 \;\Rightarrow\; \hat\lambda=3$$
$$\ell''(\lambda)=-27/\lambda^2<0 \text{ for } \lambda>0$$

1. Independence creates the product likelihood.
2. $\prod y_i!$ depends only on fixed observed data, so it may be dropped for maximisation.
3. $\exp(-9\lambda)$ contains λ and must remain.
4. Score is zero at 3; global concavity gives the unique maximum.

Answer: $\hat\lambda=3$. State the positive parameter space and why the maximum check holds throughout it.

FAILURE MODE	REPAIR IN THE ANSWER
missing independence	state iid sampling before multiplying contributions
drops $\exp(-9\lambda)$	mark every λ-dependent factor before simplifying
reports $\hat\lambda$ from memory	show score equation and admissible solution
stops at U=0	show $\ell''<0$ over $\lambda>0$ and state uniqueness

Oral self-test: What is fixed? The realised counts. What varies? λ over the positive parameter space. Why use the log? It preserves the maximiser and turns the product into a sum. What would change at a closed boundary? The endpoint becomes a candidate that must be compared.

Blank-page drill: cover the formulas and rebuild the contribution from the Poisson mass function. Say why the exponent of λ is the sample total, why the exponential term contains $n\lambda$, and why the factorial product is irrelevant only to optimisation—not to the full likelihood value.

CHECKPOINT	QUESTION TO ANSWER ALOUD
construction	which sampling assumption permits the product?
simplification	which terms vary with λ?
solution	is $\hat\lambda$ inside $\lambda>0$?
verification	is curvature negative throughout the space?

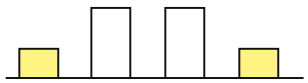
SELF-TEST → *Circle every parameter-dependent term before simplifying. "Constant-looking" is not the criterion.*

10 · Drill 2: exact sampling law

CH 2 WORKED

Given: three iid Bernoulli observations with success probability p; estimator $\hat p=\bar Y$. Find its sampling distribution and show unbiasedness.

multiplicity 1 : 3 : 3 : 1



Different ordered samples collapse to four estimator values; combinatorial coefficients survive.

$\hat p$	P($\hat P$=VALUE)	MULTIPLICITY
0	$(1-p)^3$	1
1/3	$3p(1-p)^2$	3
2/3	$3p^2(1-p)$	3
1	p^3	1

The four probabilities sum to one. Weight outputs by probabilities:

$$E[\hat p]=0\cdot(1-p)^3+(1/3)3p(1-p)^2+(2/3)3p^2(1-p)+1\cdot p^3=p$$

Therefore $\hat p$ is unbiased for every admissible p. The argument is about the random estimator over repeated samples, not the estimate from one sample.

ENUMERATION PROTOCOL

1. List possible datasets.
2. Apply the estimator to each.
3. Group datasets with the same output.
4. Add their model probabilities.
5. Check total probability = 1, then calculate moments.

PROMPT	REQUIRED RESPONSE
possible values	support of the estimator, not support of one Y_i
probability at a value	sum over all ordered samples mapped there
unbiased?	evaluate the complete expectation for general p
compare two rules	same target/model; bias, variance and MSE separately

Jensen follow-up: if the task replaces $\hat p$ by $g(\hat p)$, first decide whether g is convex or concave. Write the inequality direction before inserting $E[\hat p]=p$. Nonlinearity identifies a possible transformation bias; curvature identifies its direction.

Reason multiplicity matters: the statistical model assigns probabilities to datasets, not directly to distinct estimator outputs. Three one-success sequences and three two-success sequences produce the middle binomial coefficients.

SELF-TEST VARIATION	KEEP FIXED / CHANGE
different true p	same estimator map; dataset probabilities change

11 · Drill 3: Bernoulli bound

CH 3 WORKED

Given: n iid Bernoulli observations, success probability p. Derive information, the unbiased variance bound for p and whether $\hat p$ attains it.

one score	sample info	bound
$U_1(p)$	n / [p(1−p)]	p(1−p)/n

Information adds across independent observations; inversion gives the scalar bound for g(p)=p.

$$U_1(p)=Y/p-(1-Y)/(1-p)$$
$$I_1(p)=E[U_1(p)^2]=1/[p(1-p)]$$
$$I_n(p)=n/[p(1-p)]$$

$$\mathrm{Var}(T)\geq 1/I_n(p)=p(1-p)/n$$
$$\mathrm{Var}(\hat p)=p(1-p)/n$$

$\hat p$ attains the bound, so it is efficient among unbiased estimators under the stated regular Bernoulli model. Do not upgrade this to "best among every procedure": a biased estimator can be judged by MSE.

AUDIT LINE	BERNOULLI ANSWER
target	g(p)=p, so g′(p)=1
regular sample info	n times one-observation information
bound	inverse of $I_n(p)$
attainment	candidate variance equals bound for admissible p
scope	unbiased class under stated model

If score variance and negative expected curvature disagree, recheck support and differentiation conditions before repairing the algebra. "Efficient" is a comparison statement, so always name the class and the criterion.

12 · Drill 4: delta method

CH 4 WORKED

Given: $\sqrt{n}(S_n-\sigma^2)\to d\ N(0,2\sigma^4)$, $\sigma>0$. Find the limit for the positive square root.

$$g(x)=\sqrt{x} \longrightarrow g'(\sigma^2)=1/(2\sigma)$$

$$[1/(4\sigma^2)]\times 2\sigma^4=\sigma^2/2$$

Differentiate at the original limit σ^2, square the derivative, then multiply the original variance.

$$\sqrt{n}(\sqrt{S_n}-\sigma)\to d\ N(0,\ \sigma^2/2)$$

Answer path: transform $g(x)=\sqrt{x}$ → derivative $g'(x)=1/(2\sqrt{x})$ → evaluate at σ^2 → sandwich variance. The centre becomes $g(\sigma^2)=\sigma$.

LINE	VALUE HERE
original centre	σ^2
transformed centre	σ
squared derivative	$1/(4\sigma^2)$
original asymptotic variance	$2\sigma^4$
new asymptotic variance	$\sigma^2/2$

13 · Normal-model audit

CH 5 JUDGEMENT

Given: normal linear model with n observations and k fitted coefficients. Compare $e'e/n$ and $e'e/(n-k)$.

fitted space	⊥	residual space
rank(P)=k		rank(M)=n−k

Estimating k coefficients uses k dimensions; residual variation has n−k degrees of freedom.

$$e'e/\sigma^2\ \sim\ \chi^2_{n-k}\ \Rightarrow\ E[e'e]=(n-k)\sigma^2$$

ESTIMATOR	EXPECTATION / BIAS
$\hat\sigma^2_{ML}=e'e/n$	$((n-k)/n)\sigma^2$; bias=$-k\sigma^2/n$
$s^2=e'e/(n-k)$	σ^2; unbiased

The ML estimator is downward biased in finite samples. For fixed k, $k\sigma^2/n\to 0$, so it can remain consistent. To decide between estimators in finite samples, derive their variances and compare $\mathrm{MSE}=\mathrm{variance}+\mathrm{bias}^2$; unbiasedness alone does not settle the choice.

GEOMETRY CHECK	IDENTITY / MEANING
fitted projection	$P=P'=P^2$, rank k
residual projection	$M=M'=M^2$, rank n−k
orthogonality	$X'M=0$, so residuals are orthogonal to columns of X
independence claim	zero covariance plus joint normality

EXPLAIN → *The n−k is geometry, not a memorised correction: it is the rank of M.*

14 · Positive-scale audit

CH 6 JUDGEMENT

Given: $\sigma>0$ in a normal likelihood. Let $\lambda=\log\sigma$, hence $\sigma=\exp\lambda$. Substitute $\exp\lambda$ into ℓ and minimise $-\ell$ over unrestricted real λ.

σ>0 bounded scale	→	λ=log σ ∈ ℝ unrestricted

A one-to-one log parameterisation enforces positivity without changing the statistical model.

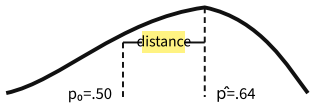
Transform back with $\hat\sigma=\exp\hat\lambda$; evaluate the original log likelihood there. Compare bounded-direct and log-scale runs on estimates, maximised objective, convergence code, final gradient and evaluations. Recommend using validity, stability and effort—not the coefficient vector alone.

NUMERICAL LAYER	QUESTION TO ANSWER
model	are parameter restrictions and likelihood contributions correct?
implementation	do analytic and finite-difference gradients agree on a small case?

15 · Drill 5: Wald vs LR

CH 7 WORKED

Given: 25 independent binary outcomes, 16 successes. Test $H_0: p=0.50$ at 5%, using χ^2_1 critical=3.841.



Same Bernoulli log likelihood, but Wald uses plug-in distance while LR uses objective drop.

UNRESTRICTED ESTIMATE

$\hat p=16/25=0.64$

WALD

$W=25(0.64-0.50)^2/[0.64(0.36)]\approx 2.13$

LIKELIHOOD RATIO

$\ell(p)=16\log p+9\log(1-p)$
$LR=2[\ell(0.64)-\ell(0.50)]\approx 1.986$
(guide: about 1.98)

Decision: both 2.13 and 1.98 are below 3.841, so do not reject H_0 at 5%. The finite-sample values differ because Wald uses a plug-in variance at $\hat p$ while LR uses the likelihood drop.

CONCLUSION MUST STATE	HERE
hypothesis + level	H_0:p=.50; 5%
statistic + reference	W≈2.13 or LR≈1.98; χ^2_1
comparison	both <3.841
decision language	do not reject; never "accept"

GENERAL REGULAR TEST

$H_0: R\theta=r,\quad q=\mathrm{rank}(R)$
$W=(R\hat\theta-r)'[R\hat V R']^{-1}(R\hat\theta-r)\to d\ \chi^2_q$

Sanity checks: the constrained estimate must satisfy the null; the unrestricted maximised objective cannot be lower; df counts independent restrictions. Near a boundary or in a small sample, Wald's plug-in quadratic approximation can be inaccurate.

Decision sentence template: "At α, the statistic is value, compared with χ^2_q critical value (or p-value). Therefore do/do not reject H_0; the sample does/does not provide sufficient evidence against the restriction." Keep the language evidential, not absolute.

PRECISION → *Keep full calculator precision until the final statistic; write the restriction count that sets df.*

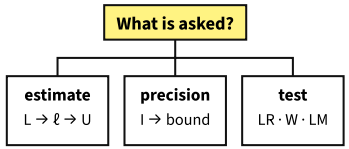
15b · Three-test chooser

FAST RECALL

IF AVAILABLE / COSTLY	USE VIEWPOINT
both constrained + unrestricted fits	LR objective drop
unrestricted fit + covariance	Wald distance
only null fit practical	Score/LM null slope

16 · If you see X → do Y

JUDGEMENT CARDS



Start with the task, then choose the object; do not begin by searching for a remembered formula.

IF YOU SEE…	DO…
"derive MLE"	model → likelihood → log → score → solve → curvature/boundary
"sampling distribution"	enumerate → map → group multiplicities → probabilities → check sum 1
"efficient"	derive information/bound; calculate candidate variance; name unbiased class
"consistent transformation"	continuous mapping
"limiting variance of $g(T_n)$"	delta method; derivative at the original limit
matrix expression	write dimensions first; check transpose and final shape
optimiser output	audit code, gradient, objective, starts, restrictions and cross-route agreement
q restrictions	reference χ^2_q; q is independent restriction count/rank

17 · Classic traps

LAST-MINUTE AUDIT

- **Likelihood:** fixed data, varying parameter; not a probability law over θ.
- **Stationary point:** score zero without maximum/boundary verification.
- **Bias:** judged from one observed estimate, or confused with consistency.
- **Information:** use equivalent formulas without checking regularity.
- **Delta:** derivative evaluated at transformed rather than original limit; derivative zero ignored.
- **Normal model:** zero covariance called independence without joint normality.
- **Optimisation:** convergence code treated as proof; log/direct objectives compared in different parameterisations.
- **Testing:** df counted as all parameters; negative LR ignored; "accept H_0" written.

18 · Exam-morning plan

COURSE STUDY ADVICE

1. Recite each method's inputs, formula, conditions and finish line.
2. Redo one likelihood derivation and one test decision from a blank page.
3. Check theorem conditions, parameter domain and matrix dimensions before algebra.