D.A.S. Fraser
Probability & Statistics : Theory & Applications

typos and fixes along with some comments and questions
by Gordon Hilton Fick

preface - v last line -> understated requirement for effective reading.. certainly you can start to appreciate the material in second year undergrad but one needs vastly more math background to make real progress with this material
'different results with different functions' -> explain...

p2 l3 Sometimes illustrations... -> should be “Sometimes illustrations in other books...”

p3 up 4 – give definition of 'partition' here. definition is in 1.3

p7 up 7 lines – 'many results in later chapters are available immediately in an invariant form' ...
provide pointers to these results? which results?

p8 problem 8 – definition of cardinality 'given' in the problem

p9 l3 'complications arise' where/when are there complications? reference?

p9 definition 1 and on -> needs more motivation... book looks to be very advanced...would readers keep with it? whole 4th year courses take on this material
..a common comment-> 'In starting to read this book, I recognized that I must read it from the beginning as the notation and terminology was not familiar to me. Then the early material looked really tough...so I lost interest'

[notation and terminology] The book is often criticized for some use of very individual notation and terminology. Is it worth going with the more 'standard' names/terms/terminology so as to remove this apparent 'impediment' for some readers?

p10 up 5 'In conclusion” ? ...of just this subsection?

p11 Cartesian product – definition

p15 l10 Borel set/class/power set – reference

p18 Exercise 18 t-statistic relative to the location 4.1 – terminology first used here...confusing?

p24 l1 upper case A should be lower case

p25 l3 reference to binomial theorem

p25 λ would be long-run average (why?) -reference

p26 10 and 11 L= and T= (or another way)

p33 l18 what is a full probability model? explain/expand

p33 l13 & l24 first use of the term 'map' and 'mapping'. reference...definition

p34 l9 add $i \to \infty$ below lim

p36 l7 counting? expand?

p36 problem 19 I_i should be I_A_i (check this... definition in Problem 22 p19...consistency?)

p36 up 10 11 $i \to \infty$ below lim (twice)

p36 problem 21 discussion at end duplicates discussion at end of section 7

p40 l8 we have no way... reference

p41... 'as an operator' – first use of the term... reference

p42 up4 'mathematical' point ... again p43 up7

p43 up11 'somewhere else'.... sentence could wait until later...tough on first read

p46 Rutherford & Geiger -> reference

p48 and p59 chi-square -> add Problem 2.2.16 Appendix II Table 3

p49 problem 8 applications last sentence "... thus voiding..." needs clarification

p50 p(.) notation first appears...why not just p?

p52 l15 S_a should be S_d
p52 'general' integral 'ordinary' integral ?

p54 up 2 present scope – reference

p55 formally

p55 section C l6 coordinates?

p57 Grummell & Dunningham -> reference

p57 l2 linear form -needs reference or expansion... discussed in second year algebra at UofT - wedge products - Grassman algebra - many/most statisticians have not seen this advanced calculus/algebra ...also p549 -proposition 2 (brevity)
linear form dy appears in some places and is absent in others -> not clear why and when?

p63 threshhold should be threshold

p68 canonical student(1) = standard student(1)=Cauchy=standard Cauchy=canonical Cauchy

p72 up 8 $(a_i, b_i)_i$ should be $(a_i, b_i]_i$

p72 up 4 y=a,b should be bold

p75 small technicality -> reference ...'as defined earlier' - where? ... 'very nice regular' --> ?

p76 E MULTIVARIATE STANDARD NORMAL DISTRIBUTION ...add STANDARD

p77 l3 z should be z_1 in the first integral and z should be z_k in the second integral

p77 l8 $z_1^2+z_1^2$ should be $z_1^2+z_2^2$

p77 l12 $z_1^2-z_2^2$ should be $z_1^2+z_2^2$

p88 increment vectors [when referred to as vectors] should be column vectors...6 sets of transposes are missing [or they should be displayed as columns]

p89 Table 2.3 should be reset/reconsidered [note the increment vectors are displayed as columns here]

p91 under independence?

p92 l7 8 where the probability.... reword/awkward

p92 determinant calculation uses determinant of a partitioned matrix... needs more explanation? easily?

p92 becomes transparent by considering Problem 2.2.15 (Gamma(k)) and Problem 3.6.10 (Dirichlet(1, 1, ... , 1) along with Problem 2.3.16

p93 l4 linear scaling?

p94 l2 to l4 uses determinant of a product is the product of determinants and uses inverse of product is product of inverses in reverse order

p94 l7 y should be bold

p94 first appearance of a kxk matrix inverse

p95 up 2 3 subscripts should be 1 and not 2

p95 up 6 sigma^2 should sigma_2 and should be aligned with the square root
...this lower triangular factorization of the matrix Σ is but one way to get to the bivariate normal

Sections 2.3 and 2.4: vectors should be column vectors throughout. The use of (z_1, z_2) without the transpose is the non-vector notation but its similarity to the corresponding row vector is 'confusing'

p97 statistics --- first use of term...rarely used except for 'order statistics' and 'rank statistics'

p105 problem 9 -algebra of symmetric functions - definition of equivalent function - equivalent first used here -> see p103 not defined
'algebraic result' --> reference

[throughout the book] exercises and problems -> essential to effective reading of this book -> referring to a problem (and its solution) in another problem

p114 up 6 lines - function has reasonable properties -> reference

p114 last line --cross-sectional? define -> many functions would work (in principle) as $\left|\frac{\partial x_1 x_2}{\partial y_1 y_2}\right|$ handles this

p119 Problem 35 is done in Section 3.6B

p119 Section 3.3 should display results for real and vector so as to be in the pattern of the rest of the book

p119 add Problems 36, 37: convolution: ratio for density on R^{k+1} ; then give development of $\boldsymbol{T}$ ~ Canonical Student(n) on R^k from z/χ ; then add Problem 5.5.11 to obtain mean and VAR for $\boldsymbol{T}$ using Lemmas 7 & 8. example of uncorrelated but dependent like 5.3.33. Add a discrete example too

p123 Section D title should be ..On The Line. (the plane and R^k should be done as well...but not in current version)

p123 Prop 6 & 7 l3 for all x and y on the plane

p125 'very important section' needed on n independent events/algebras/functions ... to give a more complete development as in sections A through C...
... then you get: $F(y_1, y_2, \ldots y_n) = \prod_i F_i(y_i)$ and then
$f(y_1, y_2, \ldots y_n) = \prod_i f_i(y_i)$ and $p(y_1, y_2, \ldots y_n) = \prod_i p_i(y_i)$
then specialize to $F_i = F$ $f_i = f$ and $p_i = p$ as in section F
Some discussion on pairwise independence not implying 'joint' independence... examples... brings the material together imho

Refer to section 3.6F for Product Spaces but independence definition is still just implied

Sections 3.3E and 3.6F should be expanded to give the various intermediate situations like 2 (or more) vectors being independent as needed to discuss
i) Example 2.4.8 which is Dirichlet independent of Gamma. Components of first vector not independent; then extended in Problem 3.6.5
ii) Problem 3.6.9 : 2 normal vectors independent; then add 3 (or more) with block diagonal matrices
iii) several (more than 2) independent as in Exercise 3.2.20, Exercise 3.3.5, Problem 4.2.7, Exercise 4.2.14

p125 IID ... appears only once...never used again

p160 proposition 4 -> awkward $P(A:C_i)$ on a separate line -> reset ...probability split

p162 3 conditions require -> equivalence ... expand? Reference?

p163 section 4.2 Take (S, A, P_C) for the plane... then set up for the distribution function $F(\boldsymbol{x}:Y=y)$ ... then go on to the probability function and density function

p173 orthogonal -- orthonormal Euclidean distance .... also 'set' on p320

p182 last line of Definition 1 – refers to (2) not (3) ... definition is not self-contained
p183 section B discrete function -- the first sentence is confusing

p194 above definition 2 ... RMS term mainly used in other contexts ... delete?

p205 Problem 20 is starred (asterisk) [label was intended to highlight harder problems but designation not used consistently] Would it be helpful to note all material requiring complex analysis

p214 Figure 5.15 Redo the scatterplots with 5 sets of random normals. Not the same set with changing correlation.

p219 lemma 3 & 4 --"marginal" mean ... "marginal" variance ...first use of the term... brief expansion worthwhile

p219 in proof of Lemma 3 – lower limit of integration off line

p219 Lemmas 3 and 4 need the existence of certain integrals. A nice example is the Canonical Student from z/χ where the degrees of freedom must be greater than 1 for the mean and greater than 2 for the variance

p230 Section 5.7 (and parts of chapter 6) All the material on characteristic functions uses results from complex analysis -- though not advanced, per se, the needed results are typically not seen until third year undergrad courses (even at UofT) or later. I think comments on the needed theory should be included and some references

p239 problem 20.... 2 asterisks? ...there are many difficult problems without asterisks. Or take out asterisks altogether.

p240 problem 28 one asterisk

p248 Theorem 3 the second part involving $m_n \text{--> } m$ is not included in the proof in section 6.3H. Do you need more conditions for $m_n \text{--> } m$ iff $F_n \text{--> } F$?

p252 Figure 6.4 -- it would be better to plot $z_1 - z$ and $z_2 - z$ to see the differences

p254 l2 definition of percentile -- more? median as a special case
l10 why do you need $p_1 < p < p_2$?
problem 17 --asterisk

p255 l3 'derived' example other than CLT?

p256 l6 this density could be included in chapter 3 and then referred to in chapter 6... it is the only density not included in chapter 3 ... and it is really interesting :)

p260 example 3, problem 6.1.11 and page 250 example 4 "continuity correction" appears in several places... due to Yates?

p270 bottom 2 paragraphs --"true" ..this term is not used again... should be?

p270 para 5 cause and effect discussions are not in chapter 12. DeLury's notes remain my favourite for these matters. go to www.ucalgary.ca/~ghfick

p272 real data examples starting with E should include the units for the data [when the data is real]

p272 E units for y [biochemical ?] again in example1 p275

p273 up1 orthogonality [see Section 2.5B]

p276 Section 7.2... since this section is about large samples, an introduction to consistency makes more sense here... could follow directly from sections of section 6.2 ....unbiasedness could wait... indeed unbiasedness is defined again in chapter 9

p276 (and other places) if the error distribution is symmetric then study of the mean is worthy... data transformations used to try to get a symmetric form... enabling the study of the mean of the transformed ... even if tests and CIs can derived from the CLT, if the distribution is asymmetric, why would one be interested in its mean?

p276 l5 first use of the term 'statistic' ...better to have a definition? is this the first use? seems late?

p277 some real 'large' samples would be better here
'small', moderate' and 'large' remain undefined... same as most nearly every book

p278 suggested, contemplated becomes hypothesized... first use... I try to keep tests of significance separate from tests of hypothesis

p279 l10 p-value = OLS now dominates... in this form, would it have been due to Fisher?
this definition is reasonable for symmetric distributions. With discrete distributions, I now use:
p-value =P(observed event or any event as likely or less likely) ... a discussion of discrepancy measures works well too

p281 l3 '...on the basis of this distribution' ... distribution could be shown here
l3 & 4 for observed data we get the observed CI.. this is a bit unclear here

p282 then the standard deviation is consistent
up 10 ...reliability used in an intuitive way here... not in a technical way
p283 paragraph 1 ... a reference to the empirical investigations? not sure what you have in mind here... standard student(n) is close to normal for n>30 or so... 'robustness' of the 't-test' is now widely quoted... maybe you are referring to our incomplete work on robustness and resistance from the late 70's

p283 l4 not section 8.3B ...nonnormal issues? not sure where this is...

p284 section G There is a lot written on these matters now... p-values now dominate.. although you will see p<0.05 implying 'significance' ...a remnant of the accept/reject ideas

p285 delete – 'for this sample size'

p285 equation (24) ...in a given ratio (the observed ratio m/n?) Wilks p275...no mention of constant ratio

p289 "medium to large" --> this brings up sample size issues... only once mentioned in the book in exercise 10.1.6... maybe more on this should be included? at least in study design material, power then has some rationale to give the researcher a guide to sample size

p295 order statistic with rank statistic link... not mentioned? order statistic first defined in exercise 1.2.14 p18

p301 (or somewhere in chapter 7) Fisher's Exact Test (FET) aught to be here... nice example of the hypergeometric... for 'small' sample 2x2 tables, FET has replaced the chi-square test... computations done easily with current software

p301 sentence before exercises -- last sentence explanation needs to be reworked... not sure of the point 'might have been obtained'

p309 l8 'very important to much of the rest of the book' ... also-> any discrete distribution... not just binomial and poisson

p310 with probability functions, the argument based on arbitrary volume is weaker. this point is often cited as a confusing section. This section needs to be strengthened so that case for the multiplicative constant is confirmed.

p311 up 9 this is the only mention of a 'relative likelihood interval' [although the name is not used] and the phrase 'with nonnormal likelihood shapes' is confusing and too early in the discussion

p312 section D Why do you need the arbitrary constant if the statistical model is discrete? : like binomial or poisson? Then discussion p316 middle does not really fix this matter

p312 just above (10) the bold capital Y should be bold lower case y

p314 first equation in Example 1 $(2\pi\sigma_0)^{n/2}$ should be $(2\pi\sigma_0^2)^{n/2}$

p316 l7 Note ... see problem 2.2.17

p320/1 section A this section needs the fact that a matrix commutes with its inverse. an nxn row orthonormal matrix is column orthonormal... not true for nxr when r is less than n

p324 l1 'convenient' ... any choice yields the same result

p332 8.4 needs more... expansion... further illustration... explicit use of 'equivalent'

p333 paragraph 3 'regularity' as discussed in section 8.1A -> not sure what was intended here?

p334 expand 'likelihood map'

p339 -fastest route- explain/expand

p339 (or somewhere near here) 'important' ...the MLE is not necessarily the likelihood statistic... example(s) (several!) would help... this is a common source of confusion

p339 exercise 8 ... the order statistic is the likelihood statistic for certain models - give some examples? Indeed, this would be the most prevalent situation when the likelihood statistic is not of fixed dimension

p341 problems 22, 23 --> both are exponential models... refer to the next section... also for the standard student one gets the the order statistic based on the squares of the t's ... and for the standard F one gets the order statistic

p345 end of section D - include references to these results [Dynkin (1951) Ferguson(1963)]
it is important to emphasize that the number of ψ functions [1 and then 2] is critical to the result.... the exponential power distributions [location] are 'curved' exponential forms

l9 ...other discrete models? basically simple? not sure... logistic regression (problem 8.4.16) ...also poisson regression

p346 eq12 o(t) needs a more careful/complete definition. This definition could be cited with each appearance of o(t) or $o(t^2)$ ...'little' oh appears in several places..each time with the same informal definition appended

p346-7 $\frac{\partial}{\partial y} l(y|.)$ "must" be one of those generated by the r ψ functions. . .. Is this argument complete? Basis vector arguments are tricky. Hoffman and Kunze addresses this matter.

equation 13 &14 have .(dot) instead of θ ... vector space of functions here ... first seen on p127 again on p572 9.6A --> I am less certain why here

there are many examples of fixed dimension likelihood statistics for certain discrete models (generalized linear models with canonical links...for example)

p347 problem 6 What is a 'reasonable' likelihood statistic?

p355 up 4 |R|<1 (why?)
just before section H – proof not included (by the methods in section F)

p356 l8 bold missing on y

p356 equation (28) the first θ should be bold $\boldsymbol{\theta}$

p357 l7 standardized

p359 prevents... expand/explain

p367 method of moments -> concern -> reference

p368 Theorem 4 -> good for nonlinear case too... least squares <-> MLE ... worth noting

p370 problem 12 scaled gamma with E(y)=1 and Var(y)= θ used in frailty models

.... also frailty test uses the 'chi-bar' distribution ... which is part discrete and part absolutely continuous.. a rare 'real' example

p370 section 9.2 second paragraph... this material does not require normal error.. worth noting

p379 up4 The mean μ ...-> unclear what this means or the point to be made... same with p380 l2 second paragraph... unbiasedness may be specific to certain models (not the only way)

p395 end of section E...'parametrization' first use of this term

p396 The Bhattacharya rth order local material works nicely here... you included this material in your graduate topics course... not too advanced though... adds another way to construct UMV... many more unbiased estimates... I have given this material as a task to the better students to try to work through

p398 up6 as discussed in Section 8.1 i.e. density means either density function or probability function

p401 up 5 lines -> expand -> develop ...reference

p401 up3 median does not have an unbiased estimate -reference?

p403 last sentence ... expand

p407 problem 23 : reference to the missing result [Uspensky(1948)] and a reference to this method of proof [there are numerous other ways to prove this result [Fraser(1951)?]
: these distributions generalize the exponential power family

p415 up7 if the null hypothesis consists...

p416/8 lemma 4/6 stated for absolutely continuous case (and preceding paragraph)..proof eludes to discrete case as analogous... does appear to be part of the statement of the lemma (critical statements on p309 provide the explanation... should include statement that integrals would be sums in the discrete case

p419 opportunity for material on sample size determinations

p421 problem 21 martingale not defined... reference... add to chapter 4?

p422 l3 a set theoretic larger ...U on its side

p423 λ probability measure over ω ... first appearance of $d\lambda(\theta)$ ... example 2 helps... probabilities for parameters first in Bayesian section 9.6
2 lines after (12) ... a set theoretic larger

p430 and 432 reference to Lehmann and Stein(1948) and King(1988)
forward reference to section E...confusing?

p433 up 6 which properties of the MVN are you referring to?
alternative definition “other authors” p432 up7

p434 3 above (6) comma n,
2 below (6) comma problems,

p435 comma example 1,

p442 bottom $4\sum(\sqrt{f_i}-\sqrt{e_i})^2$ ... sum missing

p446 mid page... the mention of the angular transformation needs a reference or expansion/motivation. Figure 10.9 gets this matter started

p450 the comment above example 3 needs expansion, clarification and/or a reference. Is it correct? Why is this comment in section G?

p451 comment at the end of example 3 just teases. it is the 'too good to be true' matter. This comment confuses.

p453 J Exercises (these are more routine than problems)

p458 Figure 11.1 $L(\mathbf{1}; y)$ should be $L(\mathbf{1}; \boldsymbol{y})$: the y should be bold; Figure 11.2 L should be $L(\mathbf{1})$

p460 l3 -> da -> more elaborate linear form matters -> expansion...justification -> compare with other references

p462 : Give the unit of measurement for serum potassium; mmol/L [then all but 2 values are elevated as normal range is 3.5 to 5.0 mmol/L] Possibly some patients with Hyperkalaemia

p462 in equation (23) $k^{-1}(\boldsymbol{d})$ should be $k_\lambda^{-1}(\boldsymbol{d})$

p462 up 4 : the term 'standard error' should not be used here.

p463 l1 λ in the range $(0,\infty)$ which goes from 'sub-Cauchy' to the normal [this change then agrees with p467 l3]

p464 up 3 : confidence interval should be probability interval

p467 Problems 10 and 11 : Give the unit of measurement for serum potassium

p480 'perhaps the model $k_\lambda(\boldsymbol{d})\,da$ may be accessible and of use' -> expand/explain -> reference ... observed $k_\lambda(\boldsymbol{d})$ and the model for possible $k_\lambda(\boldsymbol{d})$

parts of chapter 11 should be rewritten... maybe some of the new material?... references to computer programs replaced with references to software...

numerically equivalent Bayesian analyses are now available using M-H algorithms

p482 l7 chi-square-type ?

p485 reword? examining quadratic...then linear...then constant

p487 section G l8 delete comma ..the next sentence regarding bias needs expansion...bias for what?

p502 l2 [ordinary] analysis of variance table... expand...reference

p508 up 3 quadratic should linear ... or is this terminology? (A[linear] X B[linear])

p511 canonical form – define/reference

p512 Figure S.1 arrow should go from Y to a_3 as well

p515 Problem 9 Some tedious but routine... reference needed

p517 l5 continuum – define/reference

p517 Problem 17(b) I_1 needs to be checked [subscript should be $\liminf_{i\to\infty}$?]

p519 Poisson defined first on p25 in Example 6 then again p45 in Example 4, then section 2.1E

p519 part of Problem 21 up 8 – first definition of the 'general' integral

p520 l10 scalar – first use of the term

p520 l11 vector space – first time

p520 section A second paragraph last line – the linear space spanned by the columns of A
p520 section A third paragraph the 5 vectors each need a prime (') to be column vectors: (1, 1, 1, 1)'

p520 determinant result – reference

p521 Figure S.2 coordinate axes could be z's or y's. x's are used in the definition of the plane

p522 [approx =] 0.4405 radians (or 25.2394 degrees)

p523 up 11 $\lim_{n\to\infty} G_n(y)$

p525 problem 12 - “simultaneously diagonalize” needs a reference

p527 dT

p527 up 4 t referred to as 'ordinary' Student(n) but on p67 in problem 19 it is called 'standard' Student(n)... then there is the standardized Student(lambda) in Section 11.1H

p529 Problem 2 t referred to s 'ordinary' Student(n)

p529 The section should be restructured to label the problems separately from the text content. This the only section without this separation.

p529 noncentrality δ^2 should be δ ...or use same notation as for chi-square and F?

p547 $\sum$ diverges -> reference

p550 ordinary Riemann integral (first use of term 'Riemann' -Section 5.1

p552 textbooks on measure theory -reference? -> in chapter 1 references?

p559 Lemma 11 missing 'is'; should be 'is in accordance'

p576 up 7 Bayes solution... term not widely used?

p578 110 recent theoretical results--- references?

p582 section D ... I recall that the original draft (from 1972?) had a very long chapter on estimation and testing... I wish I still had all these notes... the material that survived for the book is just a taste... maybe more should go back in? The early draft (then chapter 8) included a proof that the location model that is an exponential model is either the normal or the log gamma. It would be nice to see (or reference) the proofs for the location scale and regression models. The material on invariant tests should be reinstated

p580 l8 definition of acceptance region should be given earlier in section 10.1

p581 pivotal quantities may or may not be available

p582 define an most powerful confidence region
p583 define an unbiased confidence region

p621 Poisson – should be 25 not 2,5