Errata and Comments
for
An Introduction to Econometric Theory

Andrew Petusky
(petusky@yahoo.com)

Last updated: November 25, 1999

Chapter 1

Probability

Errors

1. P. 18: Proof of DeMorgan's laws is missing superscripts. It should read

$$\begin{aligned} &\Rightarrow \quad \omega^o \notin A \text{ and } \omega^o \notin B \\ &\Rightarrow \quad \omega^o \in \tilde{A} \text{ and } \omega^o \in \tilde{B} \end{aligned}$$

2. P. 26: second line from bottom should have E_n rather than E^n.

3. P. 28: ninth line should read *disjoint sequence<u>s</u> of sets* ....

4. P. 31: line following first equation should have *for* $i = 2, 3, \ldots,$ rather than *for* $i = 1, 2, \ldots,$.

5. P. 38: the last two equations' summations should be indexed over i not t.

Comments

6. P. 24: in the last line, I think it would be clearer to write ... *all sets containing two <u>distinct</u> elements, of which....*

7. P. 27: in line 4, you have *where the sequence* $A_1, A_2, \ldots$ *ranges over all <u>disjoint</u> sequences of sets from* $\mathcal{A}$ *whose union contains* F. Do we really need *disjoint*? It seems to me we only need *disjoint* to show countable additivity. Note that in the definition of the extension of P to $\mathcal{F}$ (P. 22) you don't have *disjoint*.

8. P. 27: in the first sentence of final paragraph, you say *a place bet ... is decided ... with probability* $P(A_i) = (1/4)(3/4)^{i-1}$ .... I think it would be helpful to note that 1/4 here is from $P(win) + P(lose) = P$(the outcome of the roll is four) $+ P$(the outcome of the roll is seven) $= 3/36 + 6/36 = 1/4$ (versus 1/12 from P(the outcome of the roll is four) alone).

9. P. 28: ninth line, potentially superfluous word *disjoint*, as in [7] above.

10. P. 36: last sentence, I suggest you write as

    > Because $P_0(\Omega_0) = 1$ *and because the* $\omega_i \in \Omega_0$ *are mutually exclusive and exhaustive*, we can recover the constant of proportionality ...

11. P. 37: middle of page, you refer to Problem 12, which shows that $\mathcal{F} \cap B$ is a σ-algebra. I think you should refer to Problem 21 as well, which shows that $P(\cdot|B)$ satisfies the axioms of probability.

12. P. 39: as a final note before Section 1.7.1, you might want to note that from the ratio 244/495 it is straightforward to compute the *True Odds* (251 to 244) and also *% Casino Advantage* ($\frac{251-244}{495} \cdot 100 = 1.414$), which match the relevant entries in Table 1.1 on P. 7.

13. P. 40: in line 10 you begin a run-on sentence, which is accurate but rather confusing, especially to the uninitiated. One minor change to break it up a bit is to write

    > This result, coupled with the fact that there are only a finite number of events in the single roll σ-algebra $\mathcal{F}_P$ and therefore only a finite number of events that can cause trouble, would allow us, if desired, to modify the multiple roll probability space *so that every outcome in the sample space has the requisite behavior. This modification could be achieved, say,* by deleting from Ω_P^∞ all outcomes in the union of E_H*,* E_A*, and* $E_{A\cap H}$ (Problem 31).

14. P. 41: in first line, *However* might be changed to *Moreover*.

Chapter 2

Random Variables and Expectation

Errors

1. P. 51: second and third equations, you use F where, to be consistent with the rest of the chapter, you should use A.

2. P. 55: third equation, index of summation should be $\{i : x_i \in X(\Omega)\}$.

3. P. 56: second to last equation, indexes of summation should be $\{i : x_i \in X(\Omega)\}$ and $\{j : y_j \in Y(\Omega)\}$.

4. P. 57: third to last equation, I think it is invalid. For example, let the catch on X be 5 (so $i = 5$), let the catch on Y be 6 (so $j = 6$), and let there be four spots in common ($S = 4$). Then the formula as given is

$$f_{X,Y}(i,j) = \sum_{k=0}^{4} \frac{\binom{4}{k}\binom{4}{5-k}\binom{4}{6-k}\binom{68}{9+k}}{\binom{80}{20}}$$

which doesn't make sense unless we have negative factorials! I think changing the lower index of the summation to

$$k = \max(0, i - S, k - S)$$

will fix this.

5. P. 58-59: Some (but not all) of the summation indexes need to be changed as in [2] above.

6. P. 59: First equation should have ω as the argument, not x (compare to the third equation on P. 55):

$$X = \sum_{\{i : x_i \in X(\Omega)\}} x_i I_{F_i}(\omega).$$

7. P. 59: third line from bottom, I don't see why you have $|X_N(\omega)| \leq |X(\omega)|$ instead of $|X_N(\omega_i)| \leq |X(\omega_i)|$.

8. P. 67: Fourth equation should have ω as the argument, not x.

Comments

P. 45: You state that *The most common choice for $\mathcal{X}$ is the real line.* I suggest you give at least one example. You used Keno as an example earlier on P.45, which provides an example of a discrete $\mathcal{X}$, but not an example of a continuous $\mathcal{X}$. So it ends up being confusing

9. P. 47: I think the order of presentation could be improved. In particular, I think that you should insert Definitions 2.1 and 2.2 and Theorem 2.1 prior to the discussion about measurable functions and spaces, i.e., after the sentences *What we need for the validity of the formula... is that $\mathcal{F}_0 \subset \mathcal{F}$. A function with this property is called measurable...*

10. I don't believe that you ever define the concept of *measurable set*, although the concept is used, at least in the problem set (you define *measurable space* and *measurable function* on P. 48).

11. P. 49: following line 3, you might add something like

    > *We can solve* $X(w) = \log\left(\frac{\omega}{1-\omega}\right)$ *for* ω*:*
    >
    > $$\omega = \frac{e^X}{1+e^X}.$$
    >
    > *Note that the exponential and logarithmic functions are both increasing throughout their domains and let* $c = X(a)$ *and* $d = X(b)$*. Then* for $(c, d) \in \mathcal{X}$, the inverse image is ...

12. P. 50: it is not altogether clear that you are still talking about the coin-tossing experiment. You can improve this by modifying the sentence following the first equation:

    > Using indicators, an application of the probability function P to the interval (a, b) *for the coin-tossing experiment* can be written...

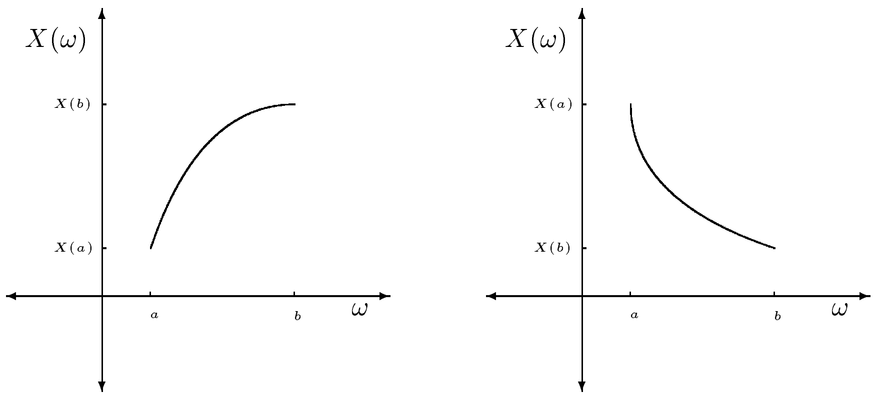
13. P. 51: final paragraph, you have been off on a bit of a digression, so I think it would be helpful to add a stronger segue, such as

    > *We return now to the derivation of the probability function for a general continuous random variable.* Consider a random variable X...

    Alternatively, come up with new sections:

2.2.1 Univariate Continuous Random Variables—a specific example
2.2.2 Indicator Functions
$\underset{\rightarrow}{NEW}$ 2.2.3 Univariate Continuous Random Variables—a general derivation
2.2.4 Bivariate Continuous Random Variables

14. P. 52 Top: it might be helpful if you added a simple picture to illustrate why the coefficients of integration are reversed.



15. P. 63/64: in section 2.4.5, to be consistent with the rest of the chapter, I recommend that you use F (and G) instead of A (and B). (Although, in contrast to [1] above, here you define $A \in \mathcal{F}$, so technically there is nothing wrong). I think this change would make the text a little easier for students just getting acquainted with the material.

16. P. 69: in Proof of Theorem 2.6, I think it would be helpful to note that the definition of conditional expectation (Definition 2.3) lets us write

$$\int_F Z(\omega)Y(\omega)dP(\omega) = \int_F \mathcal{E}(ZY|\mathcal{F}_0)(\omega)dP(\omega)$$

Chapter 3

Distributions, Transformations, and Moments

Errors

1. P. 83: you use different x-axes for the density and distribution functions, consequently they don't line up.

2. P. 84: seventh-eighth line, sentence needs extra words: *The probability of* $F \in \mathcal{F}^*$ *is the infimum* <u>*of the sum*</u> *of the probabilities of all such sets* A *that contain* F.

3. P. 85: eigth line, index of summation should be $\{i : x_i \leq x\}$

4. P. 88: fourth line should read *does not* <u>*depend*</u> *upon* x.

5. P. 94: third to last equation should have $dx = -\frac{1}{\beta}e^{-y/\beta}\underline{dy}$.

6. P. 114: last equation (marginal density) has y's and Y's in the exponent where it should be x's and X's.

7. P. 116: fourth line does not make sense (at the very least it has a redundant $g(x) + \sigma_{Y|X}$).

8. P. 118: last line of second equation should be $a'\mathrm{Var}(X)a$ rather than $a'\mathrm{Var}(X)b$.

9. P. 120/121: you have inserted X in place of Y throughout. The switch begins with the third equation from the bottom of P. 120: $\mathrm{Var}(X)$ should be $\mathrm{Var}(Y)$; in the last sentence of P. 120, *the density of* X should be *the density of* Y; the final equation of P. 120 should be $f_Y(y)$; and so on. Everywhere there is an x or X there should be a y or Y.

Comments

10. P. 81: final paragraph, you omit $H(x|n, D, N)$ (and $H(n, D, N)$) from the list of distribution functions, although you list $h(x|n, D, N)$ with the densities.

11. P. 84: first paragraph, the discussion is confusing because $X(\omega)$ maps $(\Omega^*, \mathcal{F}^*, P^*)$ into $(\mathcal{X}, \mathcal{A}, P_X)$; however, given the way you have done things up to this point, it seems that $X^*(\omega)$ might be a better choice for this random variable, *especially* since in the second paragraph you proceed to discuss $(\Omega, \mathcal{F}, P)$ but never name the random variable that takes us directly from $(\Omega, \mathcal{F}, P)$ to $(\mathcal{X}, \mathcal{A}, P_X)$. Given the development in Chapter 2, many first-years are likely to confuse $X(w)$ with the random variable that takes us from $(\Omega, \mathcal{F}, P)$ to $(\mathcal{X}, \mathcal{A}, P_X)$ (which is wrong since $X(\omega)$ maps $\omega \in \Omega^*$ from some probability space $(\Omega^*, \mathcal{F}^*, P^*)$ into $(\mathcal{X}, \mathcal{A}, P_X)$).

12. P. 85: line 5, it appears that $p = \sum p_i$ should not exceed one. If this is the case, maybe you should add this?

13. P. 85: To make the plot more clear, add $F_c(x)$ (and label $F(x)$), and include a horizontal dashed line at $(1 - p) = 0.66$.

14. P. 95/96: bottom of 95, why do you introduce $h(u, v)$ as the inverse, instead of simply $g^{-1}(u, v)$? I find the latter much clearer (especially for comparisons to the univariate case).

15. P. 100: Chi-squared is discussed here, but not listed in the index as such.

16. P. 101: I find it very helpful for the subsequent discussion to add the following after last equation

$$= \frac{1}{\sqrt{2\pi}}\sigma \int_{-\infty}^{\infty} z e^{-\frac{1}{2}z^2} dz + \frac{1}{\sqrt{2\pi}}\mu \int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2} dz$$

17. P. 104: last paragraph, I think it would be useful to add a seque here as well, something to signify that the preliminary discussion is now over, i.e.

    > *We now discuss the moment generating function.* The moment generating function, as the name suggests can be used to calculate moments...

18. P. 105, middle: you might change *which suggests that* to something like *This suggests that we take the the j^{th} derivitive w.r.t. t of the moment generating function.* This would make it clearer, emphasizing that $\frac{\partial \mathcal{E}[e^{tx}]}{\partial t} = \mathcal{E}\left[\frac{\partial e^{tx}}{\partial t}\right]$.

Chapter 4

Convergence Concepts

Errors

1. P. 128: The empirical distribution equation should have $-\infty$ rather than ∞ in the indicator function.

2. P. 131: 14 lines from the bottom, missing argument: $d \max_{1 \leq i \leq d} |X_{i,n}\underline{(\omega)} - X(\omega)|$.

3. P. 133-135: If Garland's assertion that Chapter 4 Problem 2 is incorrect, then we should have $P(|X| \geq B) \leq 2k_0 e^{-k_1 B}$.

4. P. 134: second to last equation has an e where there should be ϵ :
$$\ldots \leq 2e^{(-2n^2 \underline{\epsilon}^2/4)/4nB^2}.$$

5. P. 137: In statement of Markov's inequality, you have a strict inequality but in proof you show only weak inequality. Is this statement of Markov's inequality correct?

6. P. 140: In the second to last equation, $P(a < \ldots \leq b)$ should be $P(-a < \ldots \leq b)$ and similarly $\Phi(a)$ should be $\Phi(-a)$.

Comments

7. As a preliminary to your discussion about convergence, I suggest you insert a section on limits and on O-notation. The former topic is familiar, but probably a little hazy in most students minds, while the latter topic is probably new territory for most first-years. Hal White's <u>Asymptotic Theory for Econometricians</u> contains a nice segue from non-stochastic to random variables (P. 16); I think a similar discussion would augment your text nicely.

8. P. 129, or perhaps elsewhere: Students should be warned that $X_1, X_2, \ldots, X_n$ will sometimes refer to random samples (such as in Definition 4.1) and sometimes to statistics (such as on page 129). I find that students don't pick up on this soon enough, leading to confusion.

9. P. 134: I have a number of suggestions if you are interested.

10. P. 136: Section 4.3, the way E_n is discussed is confusing. Readers may think you are contrasting with E from the definition of almost sure convergence. To resolve this, you might define $F_n = \{\omega : |X_n - X| > \epsilon\}$ on P. 131 when you introduce a.s. convergence (and then you can rewrite your second expression for a.s. convergence as $P(\lim_{n\to\infty} F_n) = 0$ *for every* $\epsilon > 0$). Then on P. 136 you should be very explicit about what E_n is, i.e., $E_n = \{\omega : |X_n - X| > \epsilon\}$. Better yet, call this F_n too! Then when you write ... *we do not have* $P(F_n) = 0$*, as with almost sure convergence* ..., you can rest assured that you have left virtually no room for confusion. Better still, don't write $P(F_n) = 0$, but rather be explicit about the limiting process and instead write *we do not have* $P(\lim_{n\to\infty} F_n) = 0$. (This is why it would be useful to write the equation after Definition 4.3 as $\lim_{n\to\infty} P(F_n) = 0$, i.e. to facilitate the comparison with the definition of a.s. convergence.)

11. P. 136-137: probability limits are new concepts to many of the students taking this course. You might offer a simple example of a sequence of random variables that converges in probability but not almost surely.

12. P. 137: prior to Theorem 4.3, you might mention that $\mathcal{E}|X|^r$ is called *the* r^{th} *absolute moment.*

13. P. 152: top, you write *four such tests are...* but on bottom of P. 148 you list only *three....* It's tedious, but why not list all four both times?

Chapter 5

Statistical Inference

Errors

1. P. 151: is the figure correct? If W_1 and W_2 are size .05 tests, shouldn't they both have power .05 at $\theta = 0$? For the solid line this is not the case.

2. P. 154: middle, what you call a *Binomial* density looks more like a *Bernoulli* density.

3. P. 156: the section on *Strong Consistency* is numbered, but probably shouldn't be (there is no section 5.2.2.2, and in similarly structured subsequent sections, you don't number either).

4. P. 160: right-hand side of fourth equation, you have

$$-\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{d}{d\theta}\log f(x_i|\theta^0) - \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{d}{d\theta}\log f(x_i|\theta^0).$$

   This should simply be $-\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{d}{d\theta}\log f(x_i|\theta^0)$.

5. P. 160: right-hand side of third equation, you have $\overrightarrow{\mathcal{L}}$ instead of $\stackrel{\mathcal{L}}{\to}$. Is this intentional? There are numerous other examples of this (Pages 161, 166, 168, 171, 178).

6. P. 177: middle of page, the denominator of the posterior density should have p not θ:

7. P. 177: in the numerator of the last equation, you have a b where there should be β.

Comments

8. P. 149: I have some comments, but it might just be that I misunderstand. So if your interested, I'll discuss them with you.

9. P. 159/160: I think you should number the third equation on P. 159. Then before the fourth equation on P. 160, change *We now have that...* to *Making use of these results, Equation 5.X gives us....* (There are a lot of busy equations on these pages; I think this minor change would make the presentation much more readable without overly compromising the concision of the proof).

10. P. 163: You introduce the SNP density here. What does SNP stand for? Why not list it with the other densities in the appendix?

11. P. 175: You simply state Bayes rule. The derivation of Bayes rule is simple; why not add it to the relevant part of Section 1.7, or add it as an additional exercise in the Chapter 2 (or Chapter 5) problem set?

12. P. 177: The first line stinks. I suggest you change it to *Consider a beta density with parameters $y + \alpha$ and $n - y + \beta$; we have....* If you make this change, then I suggest you also remove the *Again* from the 11th line from the bottom.