Changes to the thesis "Computational methods for sums of random variables" made after the oral examination

Patrick J. Laub

October 8, 2018

1. In response to comments from Reviewer A

Addressing the major comments:

- 54_{11} **It is mentioned that component-wise fixed point iteration can be performed. Is it easy to that the equation involving the Lambert function actually defined a contraction, at least close to x^*, as would usually be required for the fixed-point iteration scheme to work?** As mentioned in my oral examination, I could not prove that the function is a contraction, though numerically it appears to behave this way on (when tested with the vector of zeros as the initial value). Upon reflection, I cannot see any advantage to using fixed-point iteration here over the Newton–Raphson scheme which is mentioned earlier in the chapter, so I have removed this discussion from the manuscript.
- Chapter 4: **Concerning the numerical experiments, it is not very clear from the context how the Laplace inversions were performed. Did you use the numerical inversion tool from Mathematica or did you code an Abate–Whitt type of programme from scratch?** It is from scratch. I added the following sentence in Section 4.5 (Numerical illustrations) to clarify: "Note, we do not use any built-in routines for the Laplace inversion, but simply implement (4.17)." Also, on pages 102–103 in Section 4.4, I:
  - removed some sentences which were repetitive,
  - removed the unnecessary header subsection 4.4.1 (but kept the content), and
  - repeated the two key equations from the introduction which give the specific inversion scheme.
- Chapter 5: **This chapter does not present the same level of clarity as the previous ones. It could be somewhat improved by being more specific about its purpose. While the methods are quite clear, the results seems somewhat mixed and confusing.** I have made many small changes which increase the clarity of the chapter and simplify the results/conclusions significantly. These include:
  - The numerical results are new and expanded, and some discussion of the patterns (related to likelihood degeneration) in these results has been added.
  - Improved wording throughout the chapter, and added more explanations.
  - Fixed error in Proposition 5.3. The cdf on the left should have been the survival function.
  - Fixed error in Example 5.6. The limit is 1 and not 2.

- Chapter 6: **. . . suffers from the same deficiency as Chapter 5 of being rather imprecise and vague in defining the problem. It would probably have helped to start with $\mathbf{129_{11}}$–$\mathbf{192_4}$ upfront, state what is the purpose and provide examples from the maximum.** I rearranged the introduction to this chapter as suggested. I removed some equations (the ones defining our estimators for the rare-maxima example) to make it very clear which problems are being addressed.

Addressing the minor comments made:

- 8^3 **". . . if we split . . . "** Replaced "if we look split the range of integration as" with "if we split the range of integration so"
- 9^{15} **"We cannot, at least, not by. . . "** Replaced "We cannot, at least, not by letting. . . " by "We cannot achieve this. Specifically, we cannot do so by letting. . . "
- 10_{10} **"The first integral. . . ", on the right hand side presumably** Added "on the right-hand side"
- 20_1 **"and will not identically"** Added "be" so it reads "will not be"
- 21^1 **"which has a stationary distribution is equal. . . "** Added "which" so it reads "which is equal"
- 28_8 **"show"** Replaced with "shown"
- 31^1 **"Nelson", should be Nelsen** Done
- 32_2 **"that we the. . . "** Added "approximate" so it reads "that we approximate the"
- 34^2 **I assume that $\phi^{(n)}$ refers to the nth derivative?** Yes, I have added this to the "Abbreviations and Notation" section of the frontmatter, under the section "Other notation"
- 65 **Formula (3.3), I suppose that λ denotes the Lebesgue measure. Why not simply writing dx?** Replaced "$d\lambda(x)$" with "dx". I simply overlooked this λ when I simplified the notation in this chapter from the earlier published version. I fixed another problem caused by the same change in notation at 68_2, where I replaced "is the given lognormal ν" with "is the lognormal with pdf $w(x)$"
- 65 **Remark 3.1. Follows by linear independence. Could be mentioned.** Added ". . . , which is a consequence of the linear independence of the orthogonal polynomials."

- 69_{14} **The step giving p_0 follows by orthogonality. If you explain the other steps, this might as well be added as well.** Added "…, and the third follows by orthogonality of the polynomials."

- 72^{10} **Mathematica, Maple, Matlab, R etc. are known to be numerically inefficient, but I suppose that the improvement was programmed in Mathematica as well?** Unfortunately one cannot look inside built-in Mathematica functions to see how they operate. Mathematica allows you to evaluate a specific Hermite polynomial, say p_k, at multiple points, say at $\{X^{[r]}, r = 1, \ldots, R\}$, so one can evaluate $\{p_0(X^{[r]}), r = 1, \ldots, R\}$ then $\{p_1(X^{[r]}), r = 1, \ldots, R\}$ and so on up to $\{p_K(X^{[r]}), r = 1, \ldots, R\}$. However I found this to be much slower than manually constructing $\{p_k(X^{[r]}), r = 1, \ldots, R\}$ by re-using the previous results $\{p_{k-1}(X^{[r]}), r = 1, \ldots, R\}$ and $\{p_{k-2}(X^{[r]}), r = 1, \ldots, R\}$ and the three-term recurrence. I concluded that Mathematica had an inefficient algorithm for evaluating $p_k(x)$ with a large k using the built-in Hermite function (perhaps it even calculates the entire sequence of $p_0(x), p_1(x), \ldots, p_k(x)$ internally then discards the intermediate results). It is not uncommon for the speed of Mathematica's functions to decrease prohibitively when the function's arguments become more extreme.

- 76 – 77 **I suppose it was on purpose that half the figures appears like chopped off at the bottom and top?** Yes it is on purpose. The motivation is we want to find to see which estimator has the smallest error. In the first subplot it is clear that the Fenton–Wilkonson and the log skew normal approximations are usually very inaccurate, so they can be ignored. The second subplots need to be zoomed in enough to be able to distinguish the remaining estimators from each other.

- 87^1 **I suppose you mean over-charging the reinsurance premium rather than the rather than the reinsurer?** I had meant that the cedant is subject to a *moral hazard* if it has a reinsurance policy based on the usual stop-loss function. This is because the cedant can exert some control of the value of S_N. Firstly it can influence the number of claims N by accepting or rejecting contracts with policyholders. Secondly it can influence the magnitude of individual claim sizes by enforcing strict interpretations of its policies and regulations, or by showing leniency and accepting every claim (perhaps hoping to gain some good-will from its customers). If the cedant's aggregated losses for a period seem like they will exceed the level a (e.g. during a large natural disaster) then they have the incentive to inflate S_N as much as possible to appear generous to their policyholders while forcing the reinsurer to pay for the excess losses. The limited stop-loss function reduces this incentive.

- 89^3 **I know what you mean, but the notation (f_N, f_U) does not make much sense.** Since this notation wasn't used again, I realised that this part of the sentence is redundant, so I have removed it. The sentence is now "We further assume that the claim sizes are independent from the claim frequency."

- 92^8 **"...we can generating the"** Replaced "we can generating the polynomials" by "we can generate the polynomials"
- 94_1 **For $r = 1$ we have the exponential distribution. Could be mentioned though obvious.** Replaced "For $r = 1$..." by "When $r = 1$ (that is, when $w(x)$ is the pdf of an exponential distribution)..."
- 124_5 **"...usually easy to implement but it takes some effort...". What is the effort and where is the problem?** Replaced

  > "The exponential tilting method is usually easy to implement but it takes some effort in this situation. Simulating each exponentially tilted Weibull variable is done via acceptance–rejection; the proposals come from the gamma distribution which is moment-matched to the asymptotic normal approximation for the exponentially tilted Weibull distribution, cf. Section 6 of [19]."

  with

  > "The exponential tilting method can be very easy to implement (in particular, when applied to distributions in the *natural exponential family*) but it takes some effort in this situation. There are no known ways to directly simulate from exponentially tilted Weibull distributions. We resort to the acceptance–rejection method with proposals coming from a gamma distribution. The specific gamma distribution is moment-matched with the asymptotic normal approximation for the exponentially tilted Weibull distribution, cf. Section 6 of [19]."

  This emphasizes the many steps it takes to simulate from the tilted distribution, in contrast to the best-case scenarios (for example, tilted exponential distributions are just exponential distributions, so the simulation effort does not increase at all).
- 128_9 **Unbiased: perhaps throw in a proof of this fact.** When I rearranged the introduction in response to the major comment on this chapter (regarding its lack of clarity) I removed this sentence. Also, after reading the new introduction and the following section, the reader can easily see that the estimators are unbiased.
- 129^6 **The second term in formula (6.3) is not well defined for $i = 1$. While of course we know what is meant, it is a bit sloppy and should at least accompanied by a comment.** Replaced "$\bigcup_{i=1}^d \{X_1 \le \gamma, \ldots, X_{i-1} \le \gamma, X_i > \gamma\}$" with "$\{X_1 > \gamma\} \cup \left(\bigcup_{i=2}^d \{X_1 \le \gamma, \ldots, X_{i-1} \le \gamma, X_i > \gamma\}\right)$". Similarly, I updated the following equation, (6.4), to avoid this $i = 1$ problem.

- 130_6 **The notation $|I| = i$ could be explained/commented on.** Added the following to the notation paragraph at the beginning of Section 6.2: "Lastly, we use the notation $\sum_{|I|=i}$ to refer to the summation over all subsets of indices $I \subset \{1, \dots, d\}$ such that I contains i indices ($|I| = i$)."

- Referring to the intersection notation:

  – 130_4 $\mathbb{P}(A_i, A_j)$ **should be** $\mathbb{P}(A_i \cap A_j)$ **(presumably).**

  – $132^{6,7}$ $\mathbb{P}(A_i, A_j)$ **again. Actually this seems to be continuing throughout, so perhaps better define it.**

  – 135^4 **You are again avoid using $\cap$, multiplying events in the indicator.**

  I added the following sentences to the notation paragraph at the start of Section 6.2:

  > "The $\cap$ notation is often dropped. We write $\mathbb{P}(A_i, A_j)$ instead of $\mathbb{P}(A_i \cap A_j)$, which is similar to the standard notation of $\mathbb{P}(X_i > \gamma, X_j > \gamma)$ to refer to $\mathbb{P}(\{X_i > \gamma\} \cap \{X_j > \gamma\})$. We also write $\mathbb{I}\{A_1 \dots A_j\}$ instead of $\mathbb{I}\{A_1 \cap \dots \cap A_j\}$."

2. In response to comments from Reviewer B

Addressing the major comments:

- Proposition 4.6 on page 97 is not correct. Reviewer B also says **"If I understand well there are consequences on the remainder of Chapter 4, but they should be easy to fix."** I have replaced this with a narrower proposition which proves the result in our situation, that is, for random sums of gamma distributed random variables. The proposition is moved back a couple of pages into the "Choice of r and m" subsection. Given the fact that the remainder of the chapter only relies on this proposition in this narrow case (of random sums of gamma variables), which is now fixed, there aren't any other changes which need to be made.

Addressing the minor comments made:

- p.92 line 9: **shouldn't the condition be $\mathbb{E}\,|X|^n < \infty$?** I have added this change. I forgot to add the absolute value signs in the exponent in the following line, so I also replaced "$\int_{\mathbb{R}} \mathrm{e}^{sx} w(x) \,\mathrm{d}x < \infty$" with "$\int_{\mathbb{R}} \mathrm{e}^{s|x|} w(x) \,\mathrm{d}x < \infty$".

- p.93 line 9: **"distributions"** Replaced "lognormal distribution" by "lognormal distributions"

- p.96: **misplace square symbol at end of proof** Fixed
- p.121, line 10: **"extracted"** Replaced "extract" with "extracted"
- p.127, line 11: **it is not clear what is the "new" distribution** When simplifying this introduction I removed this sentence.

3. Other changes

- 24_3 Replaced "$\mathbb{P}(Y > \gamma) = \frac{1}{2}$" by "$\mathbb{E}[Y] = \gamma$"
- 36_1 Removed spurious closing backet
- 39^2 Replaced "level" with "levels" so it reads "intermediate levels of rareness"
- 99^5 Replaced "The discussion in Sections 4.3 and 4.3 also apply to" with "The discussion in Sections 4.3.1 and 4.3.2 also applies to defective densities". The problem was that in the Latex these were added as subsubsections when they should have been subsections, so now they are numbered correctly.
- Cleaned up reference list. Replaced in-text citations "in [114]" with "in Chapter 2", and "in [18]" with "in Chapter 3". Remove two references which were duplicates.