Center for Quantitative Economics

Weak convergence to the *t*-distribution

Christian Schluter[†] and Mark Trede[‡]

21/2011

[†] DEFI, Aix-Marseille Université, France and University of Southampton, UK

[‡] Department of Economics, University of Münster, Germany

Weak convergence to the t-distribution*

Christian Schluter †
DEFI, Aix-Marseille Université and University of Southampton

Mark Trede‡
Westfälische Wilhelms-Universität Münster

October 7, 2011

Abstract

We present a new limit theorem for random means: if the sample size is not deterministic but has a negative binomial or geometric distribution, the limit distribution of the normalised random mean is a t-distribution with degrees of freedom depending on the shape parameter of the negative binomial distribution. Thus the limit distribution exhibits exhibits heavy tails, whereas limit laws for random sums do not achieve this unless the summands have infinite variance.

The limit law may help explain several empirical regularities. We consider two such examples: first, a simple model is used to explain why city size growth rates are approximately t-distributed. Second, a random averaging argument can account for the heavy tails of high-frequency returns. Our empirical investigations demonstrate that these predictions are borne out by the data.

*We would like to thank Krzysztof Podgórski and Gerold Alsmeyer for fruitful discussions.

†Email: C.Schluter@soton.ac.uk. DEFI, Aix-Marseille Université: 14 Avenue Jules Ferry, 13621 Aix-en-Provence, France. University of Southampton: Southampton, SO17 1BJ, UK.

‡**Corresponding author.** Email: mark.trede@uni-muenster.de. Institute for Econometrics, Center for Quantitative Economics (CQE) and Center for Nonlinear Science (CeNoS), Westfälische Wilhelms-Universität Münster, Am Stadtgraben 9, 48143 Münster, Germany, Tel. +49 251 8325006, Fax. +49 251 8322012.

1 Introduction

Limit theorems are of great interest since the limit distribution does not depend on the original sampling distribution. The most famous limit law is, of course, the central limit theorem. As early as 1809 Laplace has shown that the standardised sum of independent random variables converges in distribution to the Gaussian distribution (Stigler, 1986, p. 136ff), and since then the theorem has been considerably generalised (see e.g. Davidson (1994)).

While the sample size in these two limit theorems is non-random, sample sizes in many situations of interest are random. This has led to the discovery of limit laws for random sums. A leading case is geometric sums where the sample size is geometrically distributed, and independent of the summands. In particular, if the random summands are iid, non-negative and with finite mean, then the (non-randomly) normalised sum converges weakly to an exponential variate (Rényi (1957)). If the summands are symmetric random variables with finte variance, the limit distribution is the Laplace distribution (Kotz, Kozubowski and Podgórski (2001)). With no assumptions on the distribution of the summands, the limit is a geometric stable law (e.g. Gnedenko and Korolev (1996)), and in the case of the non-random sum a stable law (e.g. Gnedenko and Kolmogorov (1968)).

This paper considers limit laws for random means. We show that the asymptotic distribution of the random mean does *not* belong to the same distributional family as the random sum. In particular, if sample sizes are drawn from a negative binomial or a geometrical distribution, the random mean converges weakly to a t-distribution as the expected sample size goes to infinity. The number of degrees of freedom depends on the shape parameter of the negative binomial distribution, and thus equals two for the geometric distribution.

This result is of significant interest since many distributions of empirical interest exhibit heavy tails, i.e. tails that decay slowly like power functions, which lead to higher moments failing to exist. The t-distribution is, of course, heavy-tailed. Hence we have a limit law that exhibits heavy tails but which does not require the summands to have a heavy-tailed distribution. By contrast, existing limit laws for random sums do not exhibit heavy tails unless the summands themselves have infinite variance (Gnedenko and Korolev (1996)). Thus, when applied to specific contexts, the new limit law might suggest the mechanism which generates the heavy-tailed limit without presupposing it for the summands. We illustrate this in two distinct applications, one taken from urban economics, the other from finance. The first application considers the distribution of the growth rates of cities. Each city consists of economic sectors which grow randomly. Our limit law implies that growth rates of cities should be approximately distributed as a t_2 variate. We verify this prediction using data for city sizes in Germany. The second example considers high frequency stock returns and focusses on the number of transactions of a stock. The limit law implies that the one-period stock return is approximately a t_2 variate. This prediction is verified using high frequency data from the German DAX stock index.

The paper is organised as follows. In section 2 we state the limit theorem and its proof. Section 3 presents two applications. Section 4 concludes.

2 Limit distributions of random means

Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables with $E(X_k) = 0$ and finite variance $Var\,(X_k) = \sigma^2 < \infty$ for $k \in \mathbb{N}$. To keep the notation simple, we assume without loss of generality that $\sigma^2 = 1$. Let ν have a negative binomial distribution, independent of the X_i, with parameters $r > 0$ and $0 < p < 1$. The probability function is

$$P(\nu = k) = \binom{k+r-1}{k} p^r (1-p)^k \tag{1}$$

for $k = 0, 1, 2, \ldots$. This is one of the two leading cases for count models, it accommodates the overdispersion typically observed in count data (which the Poisson model cannot), and the special case of the geometric distribution ($r = 1$) is the leading case in the literature on random sums. The expectation of ν is $E(\nu) = r(1-p)/p$, so, for r given, $E(\nu) \to \infty$ as $p \to 0$. The random mean of a random number ν of draws is[1]

$$\bar{X}_p = \frac{1}{\nu} \sum_{k=1}^{\nu} X_k. \tag{2}$$

The subindex p indicates the parameter p of the negative binomial distribution. Then, by standard arguments,

$$\sqrt{\nu} \cdot \bar{X}_p \to U \sim N(0,1). \tag{3}$$

However, if the normalising factor is changed from the random variable $\sqrt{\nu}$ to the deterministic constant $\sqrt{E(\nu)} = \sqrt{r(1-p)/p}$, the limiting distribution ceases to be Gaussian. In particular, the normalised mean converges in distribution to a t-distribution with $2r$ degrees of freedom:

Theorem 1 *Under the assumptions given above, as $p \to 0$,*

$$\sqrt{\frac{r}{p}} \cdot \bar{X}_p \to T \sim t_{2r}. \tag{4}$$

Before we prove the theorem, we show the following:

Lemma 2 *As $p \to 0$, the random variates $p\nu$ and $\sqrt{\nu}\bar{X}_p$ are asymptotically independent.*

Proof of Lemma 2. Consider the joint survival function

$$\begin{aligned} P\left(p\nu > t, \sqrt{\nu}\bar{X}_p > x\right) &= P\left(\nu > \frac{t}{p}, \sqrt{\nu}\bar{X}_p > x\right) \\ &= \sum_{n > t/p} P\left(\nu = n, \sqrt{n}\bar{X}_p > x\right). \end{aligned}$$

Since ν and $\bar{X}_p$ are independent, the joint probability can be factored as

$$\sum_{n > t/p} P\left(\nu = n, \sqrt{n}\bar{X}_p > x\right) = \sum_{n > t/p} P(\nu = n) P\left(\sqrt{n}\bar{X}_p > x\right).$$

Since $n \to \infty$ as $p \to 0$, the central limit theorem applies to $\sqrt{n}\bar{X}_p$ and

$$\begin{aligned} \sum_{n > t/p} P(\nu = n) P\left(\sqrt{n}\bar{X}_p \le x\right) &\to \sum_{n > t/p} P(\nu = n) \cdot (1 - \Phi(x)) \\ &= (1 - \Phi(x)) \sum_{n > t/p} P(\nu = n) \\ &= (1 - \Phi(x)) \cdot P(p\nu > t). \end{aligned}$$

[1] In case of $\nu = 0$, the mean is defined as 0. Alternatively, one may use a different parametrisation of the negative binomial distribution with support $\{r, r+1, r+2, \ldots\}$ and expectation r/p. The following results do not depend on the parametrisation.

Hence, the joint probability can be factorized asymptotically. ■

Proof of Theorem 1. Rewrite

$$\sqrt{\frac{r}{p}} \cdot \bar{X}_p = \sqrt{\frac{2r}{2p\nu}} \cdot \sqrt{\nu}\bar{X}_p.$$

According to (3) $\sqrt{\nu}\bar{X}_p$ is asymptotically $N(0,1)$. In addition, for $p \to 0$, the random variable $2p\nu$ converges in distribution to V where $V \sim \chi^2_{2r}$. According to lemma 2, $\sqrt{2r/(2p\nu)}$ and $\sqrt{\nu}\bar{X}_p$ are asymptotically independent. Hence, as $p \to 0$

$$\sqrt{\frac{r}{p}} \cdot \bar{X}_p \to \sqrt{\frac{2r}{V}}U$$

where $U \sim N(0,1)$ and $V \sim \chi^2_{2r}$ are independent. Since $\sqrt{\frac{2r}{V}}U \sim t_{2r}$ we conclude that the normalised mean converges in distribution to a t-distribution with $2r$ degrees of freedom. ■

Corollary 3 *Let ν be geometrically distributed with parameter p. The normalised mean converges to a t-distribution with 2 degrees of freedom,*

$$\sqrt{\frac{1}{p}} \cdot \bar{X}_p \to T \sim t_2.$$

The corollary follows immediately from the observation that the geometric distribution equals the negative binomial distribution with parameter $r = 1$.

We end this Section with a few remarks. First, some of the assumptions have only been made to clarify the exposition, and can be relaxed easily. If $E(X_1) = \mu$ and $Var\,(X_1) = \sigma^2$ then $\sqrt{\nu}\left(\bar{X}_p - \mu\right)/\sigma \to U \sim N(0,1)$ and

$$\sqrt{\frac{r}{p}} \cdot \frac{\bar{X}_p - \mu}{\sigma} \to T \sim t_{2r}. \tag{5}$$

Second, it is not necessary to assume that $X_1, X_2, \ldots$ are i.i.d. Any version of the central limit theorem with non-identical and/or dependent observations can be applied as long as the sample size is independent of the $X_1, X_2, \ldots$. Third, the distributional assumptions concerning the random sample size can also be relaxed. As long as $2\nu p$ converges weakly to a χ^2-distribution as $p \to 0$, the results still hold. In particular, ν could be a shifted negative binomial or geometric distribution. Fourth, the theorem can also be transferred to sample statistics other than the mean. If a sample statistic can be estimated by a$\sqrt{n}$-consistent estimator (e.g. quantiles), the theorem continues to hold.

3 Applications

We present two distinct applications, one taken from urban economics, the other from finance, in order to demonstrate how the weak convergence theorem can be applied to explain striking empirical regularities. In particular, both example show how heavy-tailed distributions can emerge, and we demonstrate that the theoretical prediction is born out by the data.

3.1 The growth rates of cities

The size distribution of cities is often found to be of the Pareto type, an empirical regularity known as Zipf's law (see e.g. Gabaix (1999) or Córdoba (2008)). Accordingly, assume that the size X_i of city i, measured as the number of inhabitants, follows a Pareto distribution with scale parameter $x_0 > 0$ and shape parameter $\alpha > 0$,

$$P(X_i > x) = \begin{cases} x_0^{\alpha} x^{-\alpha} & \text{for } x > x_0 \\ 0 & \text{for } x \leq x_0 \end{cases} \tag{6}$$

Next, assume that there are S_i economic sectors in city i and that the number of sectors S_i depends on the city size in the following way

$$S_i = C + \lambda \ln X_i. \tag{7}$$

A simple model consistent with (7) is Christaller's central place theory Christaller (1966), and empirical evidence in support of (7) is reported in (Mori et al., 2008, Figure 6). Equations (6) and (7) imply that S_i has a shifted exponential distribution with cdf

$$P(S_i \leq s) = \begin{cases} 1 - B\exp\left(-\frac{\alpha}{\lambda}s\right) & \text{for } s > C + \lambda \ln x_0 \\ 0 & \text{else} \end{cases}$$

where $B = x_0^{\alpha} e^{\alpha C/\lambda}$.

The discrete counterpart of the (shifted) exponential distribution is a (shifted) geometrical distribution. Hence, the limit theorem can be used to explain the unconditional growth rate distribution of city sizes. In particular, if sector j grows at the random rate r_j with $E(r_j) = \mu$ and $Var(r_j) = \sigma^2 < \infty$, then the average growth rate R_i of city i is the random mean

$$R_i = \frac{1}{S_i}\sum_{j=1}^{S_i} r_j.$$

According to Theorem 1,

$$\left(p\sigma^2\right)^{-1/2}(R_i - \mu) \sim t_2, \tag{8}$$

i.e. the city size growth rate distribution should follow approximately the t_2-distribution.

The exact Pareto assumption of equation (6) has been made mainly for expositional simplicity. This could be generalised to a domain of attraction assumption, so that (6) holds for sufficiently large x, and x_0 can be replace by any slowly varying function. Thus we can accommodate weaker versions of Zipf's law (see also Schluter and Trede (2011) for further discussions).

3.1.1 Empirical investigation

We consider the distribution of city growth rates using data for German cities (or, more precisely, "Gemeinde"). The data are provided by the German Federal Statistical Office, and very accurate given the legal obligation of citizens to register with their respective town halls. For a detailed description of the data, see Schluter and Trede (2011). In this illustration, we select two representative years, 1995 and 1996, in which the number of cities is $n = 14\,551$. Our earlier work confirms that the upper right tail of the city size distribution is indeed heavy and thus of the Pareto type. The estimates of shape parameter α (or, more precisely, of the extremal index) range between 1.24 and 1.31 depending on the estimation method, and did not change much over the period under investigation, 1995-2006.

Turning from the size data to the annual growth rate of cities, the mean growth rate is 1.02%. The proportion of observations with small absolute values is large, the 0.05-quantile is -3.33% and the 0.95-quantile is 6.4%. However, the range is large with a smallest growth rate of -57% and a largest one of $+100\%$. Figure 1 shows the histogram of the growth rates.

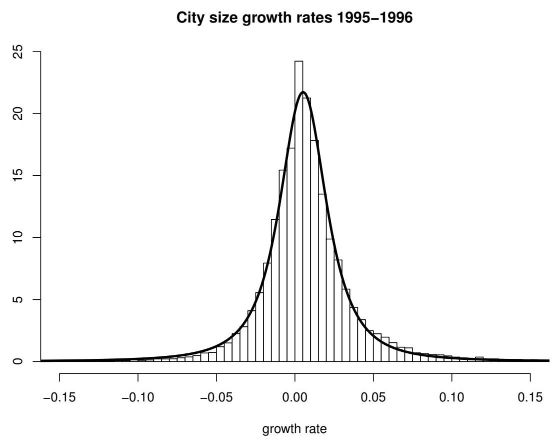

Figure 1: Histogram of the city size growth rates and a fitted t-distribution with two degrees of freedom.

If the preceding hypotheses are met, growth rate are distributed as a t_2 variate by (8). We depict the density in Figure 1, having estimated the location parameter μ and scale parameter σ in equation (8) by maximum likelihood, the estimates being $\hat{\mu} = 0.00521$ (standard error $se = 0.000175$) and $\hat{\sigma} = 0.01626$ ($se = 0.000190$). It is evident that the limit law fits the actual data closely.

If the degrees of freedom parameter is estimated along the location and scale parameters, the maximum likelihood estimated degrees of freedom is $\widehat{df} = 2.005$ with a standard error of 0.043. Hence, the hypothesis that the number of degrees of freedom equals 2 cannot be rejected at usual significance levels.

We conclude by noting one implication of these results for estimation and inference. Averaging a random number of sectoral growth rates with finite variances results, in the limit, in the overall growth rate distribution having infinite variance. This implies that statistical procedures relying on finite second moments are invalid.

3.2 High-frequency stock returns

Stock returns distributions are well-known to exhibit heavy tails (see the discussion in e.g. Schluter and Trede (2008)). We proceed to propose a transaction-based model which, in conjunction with our limit theorem, yields a heavy-tailed returns distribution.

Let $K_{t,0}$ denote a stock price at the begin to period $t = 1, 2, \ldots$. Suppose that there is a random number ν_t of market transactions during period t. For simplicity the transactions are equally spaced over period t, and the random length of the sub-periods is $\Delta_t = 1/\nu_t$. Let each transaction cause a "periodised" return $r_{t,i}$ where $r_{t,i}$ is i.i.d. with $E(r_{t,i}) = 0$ and $Var(r_{t,i}) = \sigma^2 < \infty$, $i = 1, \ldots, \nu_t$. The stock price after the jth transaction in period t is $K_{t,j} = K_{t,0} \cdot \exp\left(\sum_{i=1}^{j} r_{t,i}\Delta_t\right)$, hence

$$K_{t+1,0} = K_{t,0} \cdot \exp\left(\sum_{i=1}^{\nu_t} r_{t,i}\Delta_t\right).$$

Note that the number of transactions ν_t has an impact on the return distribution of each single transaction. The more transactions occur, the lower their individual variances. This may be thought of as an liquidity effect, as more transactions tend to make the price movements smoother. Given the number of transactions, the variance over one entire period is always σ^2.

The unconditional distribution of the one-period returns

$$\begin{aligned} R_{t+1} &= \ln\left(K_{t+1,0}/K_{t,0}\right) \\ &= \frac{1}{\nu_t}\sum_{i=1}^{\nu_t} r_{t,i} \end{aligned}$$

is a random mean of a random number of returns. If the number of transactions ν_t has a negative binomial distribution with shape parameter r (or geometric distribution, $r = 1$), then

$$\left(p\sigma^2\right)^{-1/2}(R_i - \mu) \sim t_{2r}, \tag{9}$$

as $p \to 0$. The limit distribution is thus heavy-tailed even though $Var(r_{t,i}) = \sigma^2 < \infty$; however, the distribution has a finite variance if the degrees of freedom exceed 2.

3.2.1 Empirical investigation

We consider high frequency returns of the German stock index DAX. The data are provided by the EUREX database, the sample includes 2,612,189 observations (at the 1 second sampling frequency) from 2nd January until 28th April 2006, observed for 90 trading days over 18 weeks. The daily trading phase starts at 9 am and ends at 5.45 pm. In order to prevent noise due to the market's micro structure we compute 5 min returns. Overnight returns are discarded, as are 84 zero returns. The remaining number of observations is 8,541. Figure 2 depicts the histogram of the 5 min returns.

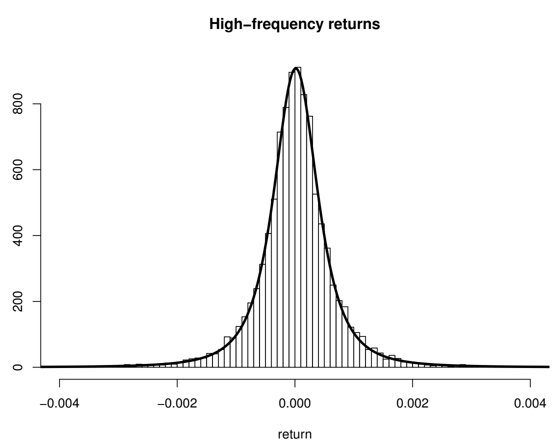

Figure 2: Histogram of the 5 minute returns of the DAX index and density of a fitted t-distribution

The figure also depicts the fitted density of a scaled t_{2r}-distribution, based on equation (9), with location and scale parameters (μ and σ) as well as the degrees of freedom (r) all estimated by maximum likelihood. The estimates are $\hat{\mu} = 1.4146 \times 10^{-5}$ (with standard error $se = 0.8498 \times 10^{-5}$), $\hat{\sigma} = 3.9716 \times 10^{-4}$ ($se = 0.1658 \times 10^{-4}$), and $\hat{r} = 2.4090$

($se = 0.0584$). The limit law fits the actual data very closely. We conclude by observing that the estimated degrees of freedom are significantly larger than 2, indicating that although the return distribution is heavy-tailed its variance is finite.

4 Conclusion

If the sample size is not deterministic but has a negative binomial or geometric distribution, the limit distribution of the normalised random mean is a t-distribution with a number of degrees of freedom depending on the shape parameter of the negative binomial distribution. Hence the limit distributions of the random mean and the random sum belong to two different families.

The limit distribution exhibits exhibits heavy tails without presupposing it for the summands, and may thus help explain several empirical regularities. We have considered two such examples: first, a simple model is used to explain why city size growth rates are approximately t-distributed. Second, a random averaging argument can account for the heavy tails of high-frequency returns.

References

Christaller, W. (1966). *Central Places in Southern Germany.* Englewood Cliffs: Prentice Hall.

Córdoba, J. C. (2008). A generalized Gibrat's law. *International Economic Review 49*(4), 1463–1468.

Davidson, J. (1994). *Stochastic Limit Theory.* Oxford: Oxford University Press.

Gabaix, X. (1999). Zipf's law for cities: An explanation. *Quarterly Journal of Economics 114*(3), 739–767.

Gnedenko, B. and A. Kolmogorov (1968). *Limit Distributions for Sums of Independent Random Variables.* Reading: Addison-Wesley.

Gnedenko, B. V. and V. Y. Korolev (1996). *Random Summation: Limit Theorems and Applications.* Boca Raton: CRC Press.

Kotz, S., T. J. Kozubowski, and K. Podgórski (2001). *The Laplace Distribution and Generalizations.* Boston: Birkhäuser.

Mori, T., K. Nishikimi, and T. E. Smith (2008). The number-average size rule: A new empirical relationship between industrial location and city size. *Journal of Regional Science 48*(1), 165–211.

Rényi, A. (1957). On the asymptotic distribution of the sum of a random number of independent random variables. *Acta Mathematica Hungarica 8*, 193–199.

Schluter, C. and M. Trede (2008). Identifying multiple outliers in heavy-tailed distributions with an application to market crashes. *Journal of Empirical Finance 15*, 700–713.

Schluter, C. and M. Trede (2011). Gibrat, Zipf, Fisher and Tippett: City size distributions reconsidered. Working Paper.

Stigler, S. M. (1986). *The History of Statistics: The Measurement of Uncertainty Before 1900.* Cambridge: Harvard University Press.