Working Paper Sustainability and Innovation
No. S 1/2014

Patrick Plötz

How to estimate the probability of rare long-distance trips

Abstract

The vehicle distances travelled by individual users can very strongly between different days. This is particularly problematic for electric vehicles since trips larger than the electric range clearly reduce the vehicle's utility. Here we estimate the number of days with driving distance larger than a given threshold for individual users based on their observed driving behaviour. The general formalism is developed and estimates for the main observable and standard errors are derived based on the assumption of individual log-normal distributed daily vehicle kilometres travelled. Numerical simulations of driving profiles demonstrate the validity and accuracy of the analytical results.

Contents

1 Introduction 1

2 Derivation of Estimates 2
2.1 Days per Year Requiring Adaptation 2
2.2 Standard Error Estimate 3
2.3 Annual Vehicle Kilometres Travelled 4

3 Numerical Results 5
3.1 Days per Year Requiring Adaptation 5
3.2 Annual Vehicle Kilometres Travelled 7

4 Summary 8

1 Introduction

Electric vehicles (EVs) charged with renewable electricity are a possible way of reducing green house gas emissions from the transport sector without sacrifice of individual car-based mobility [1, 2]. But the limited electric driving range of electric vehicles is a major hurdle for many consumers. Given a certain range, it is natural to ask what share of users can cover all their trips in one day. However, the underlying data is often cross-sectional with many users but only one or very few days of observation per user. For realistic estimates of the required range for EVs, the frequency of long-distance trips for each individual user is needed and does not necessarily coincide with that of many users. First results include the frequency of long-distance trips of a limited sample over longer time periods [3, 4], but in most cases the observed driving period is rather short, between one day and a few weeks. Furthermore, a general framework to measure and compare the occurrence of long-distance trips for many individual users is still lacking.

The statistical distribution of daily vehicle kilometres (VKT) travelled with applications to EVs has drawn attention in the literature and found new stimulus by the market introduction of EVs and GPS based measurements [3, 4, 5, 6, 8, 7]. However, the analysis of the distribution of individual users with long observation periods is still a young field [5, p. 218] and several distributions have been proposed. There seems to be overall agreement that both individual and cross-sectional VKT distributions are peaked and right skewed. Greene [3] and Lin et al. [6] discuss three distributions that have the aforementioned shape: the Weibull, log-normal and Gamma distribution. That analyse two data sets and argue that the Gamma distribution is most suitable [3, 6]. However, Blum [7] compares data from one year of driving in Canada of 76 vehicles and concludes that the log-normal distribution provides the best fit both for the most drivers and for the largest number of total trips. This is consistent with the the log-normal distribution as best fit to cross-sectional daily VKT [12]. Furthermore, the log-normal has the advantage of simplicity in analytical calculations compared to the gamma distribution. So far, the numerical results on the best fitting two-parameter distribution for daily VKT is not conclusive, but in the following, we will use log-normal distribution to describe the distribution of individual daily VKT.

The aim of the present paper is to develop a methodology for assessing the frequency of long-distance trips for individual users extrapolated from a limited observation period. The method is based on an assumed probability distribution for individual daily driving distances. We use a log-normal distribution but many results are more general. The outline is as follows. The formalism and theoretical results are presented in section 2, they are illustrated by and compared to numerical results in section 3, followed by a summary in section 4.

2 Derivation of Estimates

2.1 Days per Year Requiring Adaptation

Let us assume that the daily driving distances r of an individual user i are log-normal distributed, i.e.

$$f_{\mu_i \eta_i}(r) = \frac{1}{r\sqrt{2\pi}\eta_i} \exp\left[-\left(\frac{\ln r - \mu_i}{\sqrt{2}\eta_i}\right)^2\right]. \tag{1}$$

We want to calculate the number of days $D(L)$ per year on which the driving distance r is larger than a threshold L. In other words, $D(L)$ are the number of days per year that would require adaptation from a BEV user. This approach can then be applied to each user i with his μ_i and η_i individually, but the derivation is of course general and we will therefore suppress the index i in the following.

If we measure the daily driving distances r_l over $l = 1, \ldots, n$ days out of N total days of observation, we can estimate the parameters μ and η of the individual user's driving distribution from the sample mean M and sample variance S as follows (estimates for the underlying parameters are denoted by hats ˆ and their numerical values by latin letters, see table 1 for an overview)

$$\hat{\mu} = M \equiv \frac{1}{n}\sum_{l=1}^{n} \ln r_l \quad \text{and} \quad \hat{\eta}^2 = S^2 \equiv \frac{1}{n-1}\sum_{l=1}^{n}(\ln r_l - \hat{\mu})^2. \tag{2}$$

Alternatively, one could use $m = \frac{1}{n}\sum_l r_l$ and $v = \frac{1}{n-1}\sum_l (r_l - m)^2$ to obtain $\hat{\mu} = \ln(m^2/\sqrt{v + m^2})$ and $\hat{\eta} = \sqrt{\ln(1 + v/m^2)}$ (see [9, 10] for a discussion and comparison).

The probability $\Pr(L > r)$ for driving a trip long than L is then given by

$$\Pr(L > r) = \int_L^\infty f_{\mu\eta}(r)\mathrm{d}r = 1 - \int_0^L f_{\mu\eta}(r)\mathrm{d}r = 1 - F_{\mu\eta}(L) \tag{3}$$

where $F_{\mu\eta}(x) = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{\ln x - \mu}{\eta\sqrt{2}}\right)\right] \approx \left[1 + (\frac{\mathrm{e}^\mu}{x})^{\pi/(\eta\sqrt{3})}\right]^{-1}$ is the cumulative distribution function (cdf) of the log-normal distribution.

If the user has been driving on n days out of the total N days of observation, we estimate the fraction α of days the user is driving by $\hat{\alpha} = a = n/N$ and obtain the number of days D per year with a daily driving of more than L kilometres as

$$\begin{aligned} D(L) &= 365\,\alpha \Pr(L > r) = 365\,a\,(1 - F_{\mu\eta}(L)) \\ &= 365\,a\left[\frac{1}{2} - \operatorname{erf}\left(\frac{\ln L - \mu}{\eta\sqrt{2}}\right)\right] \approx \frac{365\,a}{1 + (\frac{L}{\mathrm{e}^\mu})^{\pi/(\eta\sqrt{3})}}. \end{aligned} \tag{4}$$

This is the desired result. Based on a finite number of driving and observation days, one estimates the number of days $D(L)$ with vehicle kilometres travelled larger than a given threshold L. Furthermore, the general dependence is given by $D(L) = 365\,a/(1 + c_1 L^{c_2})$ with user dependent constants $c_1 = \mathrm{e}^{-\mu c_2}$, $c_2 = \pi/(\eta\sqrt{3})$. This is clearly seen in the numerical results of figure 1 below.

One could, of course, simply estimate this same quantity $D(L)$ by multiplying the fraction of days d/N for which $r_l > L$ by 365, but this estimate changes abruptly with slight changes in L or a few more driving days. Note that the first line of eq. (4) is independent of the assumed distribution function and holds for the Gamma distribution as well. For example, if the measured vehicle was observed for 14 days, moved on eleven days and drove more than 150 km on two days, one would estimate $D \approx 365 \cdot 11/14 \cdot 2/11 \approx 52$. For a vehicle with the same characteristics but three out of eleven days with $r_l > 150$ km one obtains the 50% higher $D \approx 78$. This reasoning ignores the information given by all other driving days and would lead to a simple step function in figure 1 below. The result in eq. (4) has the advantage of being a continuous estimate in L despite the limited observation time and uses the information given by all other driving days, too. Of course, the assumption of a log-normal distribution has to be made and its parameters have to be estimated from a finite sample. However, standard statistical methods are available for this (see next section).

2.2 Standard Error Estimate

Since we do not know α, μ and η, they have to be estimated from the sample, i.e. from each user's individual driving profile r_l, as $\hat{\alpha}, \hat{\mu}$, and $\hat{\eta}$. Using eq. (4), one obtains an estimate for $\hat{D}(L)$. Furthermore, an estimate of the standard errors associated with these estimate is helpful to understand the precision of (and confidence intervals for) the estimated number of days per year requiring adaptation. Assuming small standard errors and independence of the variations, i.e. $\mathrm{cov}(\mu, \eta) = 0 = \mathrm{cov}(\mu, \alpha)$, we can propagate the errors [11] to obtain

$$\hat{\sigma}^2_{\hat{D}} = \left(\frac{\partial D}{\partial \alpha}\right)^2 \hat{\sigma}^2_{\hat{\alpha}} + \left(\frac{\partial D}{\partial \mu}\right)^2 \hat{\sigma}^2_{\hat{\mu}} + \left(\frac{\partial D}{\partial \eta}\right)^2 \hat{\sigma}^2_{\hat{\eta}}. \tag{5}$$

The partial derivatives are obtained straightforwardly

$$\begin{aligned}
\frac{\partial D}{\partial \alpha} &= D/\alpha = 365 \cdot (1 - F_{\mu,\eta}(L)) \\
\frac{\partial D}{\partial \mu} &= 365\, a\, L \cdot f_{\mu\eta}(L) \\
\frac{\partial D}{\partial \eta} &= 365\, a\, L \cdot f_{\mu\eta}(L) \left(\frac{\ln L - \mu}{\eta}\right).
\end{aligned}$$

Parameter	Estimator	variance estimator
driving day fraction α	$\hat{\alpha} = a = n/N$	$\hat{\sigma}^2_{\hat{\alpha}} = a(1-a)/N$
scale μ	$\hat{\mu} = M$	$\hat{\sigma}^2_{\hat{\mu}} = S^2/n$
shape η	$\hat{\eta} = S$	$\hat{\sigma}^2_{\hat{\eta}} = S^2/(2n)$

Table 1: Overview of parameter estimators and standard error estimators. We use the sample mean $M = \frac{1}{n}\sum_l \ln r_l$ the sample variance $S^2 = \frac{1}{n-1}\sum_l(\ln r_l - M)^2$. The number of days observed is given by N and the number of days on which the vehicle has been driven is $n \leq N$.

The standard error for the number of days requiring adaptation thus reads

$$\hat{\sigma}^2_{\hat{D}} = \frac{365^2\alpha}{N}\left[(1-F_{\mu\eta})^2(1-\alpha) + (Lf_{\mu\eta}S)^2\left(1+\frac{\ln L - \mu}{\sqrt{2}S}\right)\right]. \qquad (6)$$

The half width of a confidence interval around $D(L)$ with confidence level τ is then approximately given by $z_{1-\tau/2}\cdot\hat{\sigma}_{\hat{D}} \propto 1/\sqrt{N}$ with relatively slow convergence as the square root of the observation period N.

2.3 Annual Vehicle Kilometres Travelled

In a similar fashion, one can use the average daily driving distance $\bar{r} = \frac{1}{n}\sum_l r_l = \exp[\mu + \eta^2/2]$ to estimate the annual vehicle kilometres travelled (VKT) $R = 365\,\alpha\,\bar{r} = 365\,\alpha\exp[\mu+\eta^2/2]$ and its error. Performing similar steps as above we obtain

$$\begin{aligned}\hat{\sigma}^2_{\hat{R}} &= \left(\frac{\partial R}{\partial \alpha}\right)^2\hat{\sigma}^2_{\hat{\alpha}} + \left(\frac{\partial R}{\partial \mu}\right)^2\hat{\sigma}^2_{\hat{\mu}} + \left(\frac{\partial R}{\partial \eta}\right)^2\hat{\sigma}^2_{\hat{\eta}} \\ &= \frac{R^2}{aN}\left(1-a+S^2+S^4/2\right). \qquad (7)\end{aligned}$$

Accordingly, the relative error reads $\hat{\sigma}_{\hat{R}}/R = \frac{1}{\sqrt{aN}}\left(1-a+S^2+S^4/2\right)^{1/2} \propto \frac{1}{\sqrt{N}}$. As before, the convergence is quite slow and very long observation period are required if high precision is aimed at.

3 Numerical Results

3.1 Days per Year Requiring Adaptation

We performed a Monte Carlo simulation to assess the quality of our estimates. Figure 1 shows the results of 1,000 runs with daily VKT generated from a log-normal distribution with $\mu = 3.43$ and $\eta = 1.20$ for the 'average' German driver [12]. The observation period is $N = 14$ days and the number of days with driving n is drawn from a binomial distribution in each run with $\alpha = 5/7$. For each randomly generated driving profile, the estimated number of days requiring adaptation has been estimated according to eq. (4). Since the driving profiles are randomly generated, many different number of days are estimated. Figure 1 shows the mean and median as well as the upper and lower quartiles of the distribution of estimated days D. Also shown is the "real" number of days as calculated with the correct values $\mu = 3.43$, $\eta = 1.20$ and $\alpha = 5/7$.

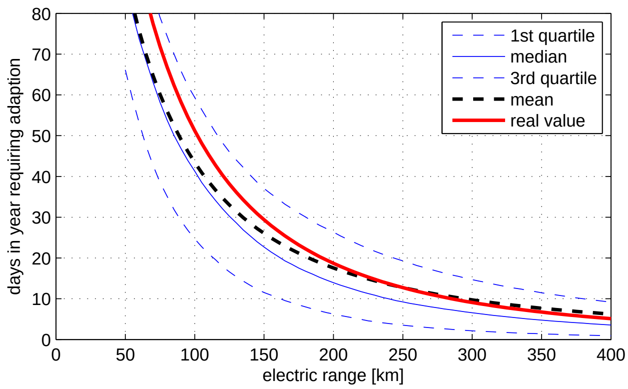

Figure 1: Monte Carlo simulation for the estimated number of days per year requiring adaption for different electric driving ranges.

Similar to the simulation described above, we generated 1000 driving profiles from fixed numerical parameter with one to 52 weeks of observation. Figure 2 shows the distribution of the estimated number of days D from one week of observation and from a whole year of observation. The distribution for a longer observation period is clearly much more peaked since much more daily driving distances are contained in the simulated data and a much better estimate of the original values of μ, η and α and accordingly a more precise estimate of D are possible. The right panel of figure 2 shows the numerical and analytical results for the full width of a 68% confidence interval, i.e. $2\hat{\sigma}_D$ for the analytical curve and the difference between the 0.841- and 0.159-quantile of the numerical distribution of D for a given observation period. The agreement between the simulated and calculated standard errors is quite remarkable.

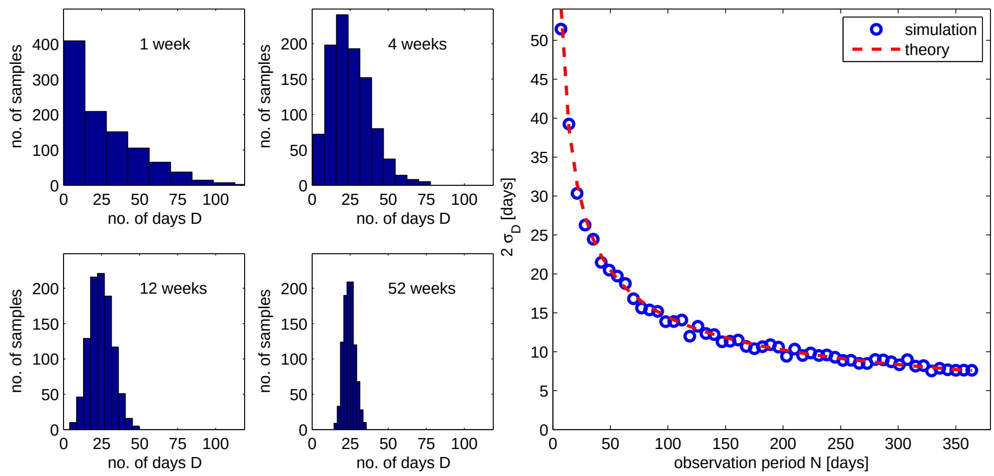

Figure 2: Monte Carlo simulation for the number of days requiring adaptation for different observation periods. Each data point in the right panel corresponds to the width of the distribution on the left of 1,000 randomly generated driving profiles.

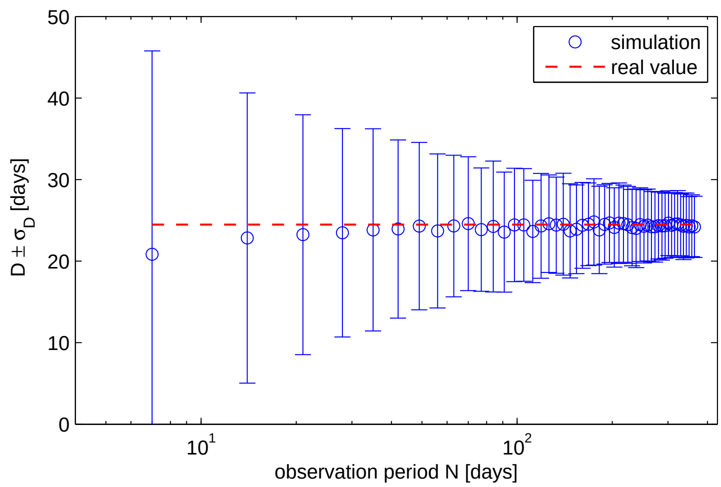

Figure 3: Estimated number of days requiring adaptation $\hat{D}$ with 68% confidence bands from the numerical simulation. Show is also the 'real' value from the underlying parameters.

Finally, we can use the standard error $\hat{\sigma}_D$ to add confidence bands the estimate for the number of days requiring adaptation $D(L)$ from eq. (4). This is shown in figure 3. Since the samples for the simulation are quite large (1,000 runs), the numerical average is very early close to the real value. However, when using a real driving profile, the underlying sample size to estimate an individual value for

D is only one and we have to use the full error bars to account for the underlying uncertainty of the estimate. Please note that the width of the 68% confidence interval is smaller than the estimated number of days, i.e. the signal to noise ratio is smaller than unity, after five to six weeks.

3.2 Annual Vehicle Kilometres Travelled

Using the same parameters as in the previous section, i.e. $\mu = 3.43$, $\sigma = 1.20$ and $\alpha = 5/7$, we simulated 1,000 driving profiles for fixed observation period and estimated R for each driving profile. The distribution of estimated Rs for fixed observation period is used to obtain a numerical $\hat{R}$ and the difference between the 0.841- and 0.159-quantile is used as numerical approximation to $2\sigma_R$. These are then compared to the "theoretical" values $\hat{R} = 365\,\hat{\alpha}\exp[\hat{\mu} + \hat{\sigma}^2/2]$ and eq. (7). We obtain $R \approx 16,500\text{km}$, $\sigma_R \approx 32,500\text{km}/\sqrt{N}$ and $\sigma_R/R \approx 1.97/\sqrt{N}$. Numerical examples are stated in table 2.

N [days]	7	14	28	63	91	182	364
$\hat{\sigma}_R$ [km]	12,290	8,690	6,150	4,100	3,400	2,400	1,700
$\hat{\sigma}_R/\hat{R}$	74.3%	52.6%	37.2%	24.8%	20.6%	14.6%	10.3%

Table 2: Numerical example for the precision of the estimated annual VKT in terms of the standard error $\hat{\sigma}_R$ and relative error $\hat{\sigma}_{\hat{R}}/R$ according to eq. (7).

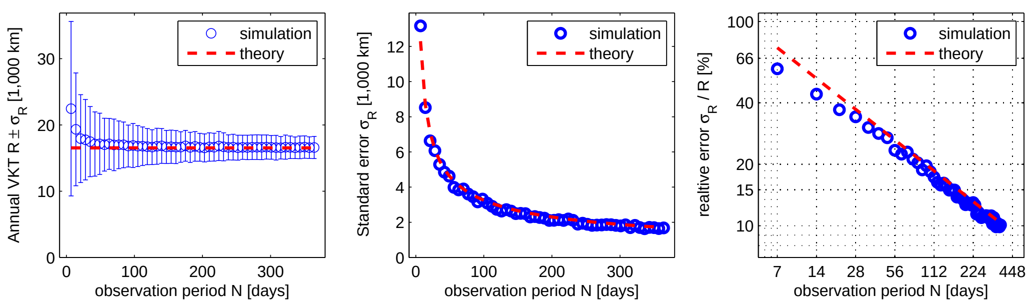

Figure 4: Comparison between estimated and simulated annual VKT for increasing observation period. Shown are the estimated annual VKT with 68% confidence interval (left panel), the simulated and estimated standard error σ_R (middle panel) and the relative error σ_R/R (right panel, note the logarithmic scales).

The numerical and analytical results are compared in figure 4. We observe good overall agreement between the simulation results and the analytical estimates. However, the left panel of figure 4 indicates that annual VKT are systematically overestimated for short observation periods (up to five weeks), we observed similar behaviour in different numerical runs.

4 Summary

We derived an estimate and standard error for the number of days with vehicles kilometres travelled larger than a given threshold based on the assumption of log-normal distributed daily VKT. The numerical results show the usefulness and precision of the calculated estimates and precision. The precision of the estimated driving range and number of days requiring adaptation increases quite slowly with inverse square root of the number of observation days. If the observation period is longer than four weeks, the width of an estimated 68% confidence interval is smaller than the estimated number of days. Future research should compare the results derived here to real world driving data as in [7] and find similar estimates for gamma distributed daily VKT.

References

[1] C.C. Chan: The state of the art of electric, hybrid, and fuel cell vehicles, *Proceedings of the IEEE* Vol. 95 (2007), 704.

[2] C.E. Thomas: How green are electric vehicles?, *International Journal of Hydrogen Energy* vol. 37 (2012), 6053-6062.

[3] D.L. Greene: Estimating daily vehicle usage distributions and the implications for limited-range vehicles, *Transportation Research Part B: Methodological* 19 (1985), 347–358.

[4] N.S. Pearre, W. Kempton, R.L. Guensler, V.V. Elango: Electric vehicles: How much range is required for a day's driving, *Transportation Research Part C: Emerging Technologies* 19 (2011), 1171-1184.

[5] R. Smith, S. Shahidinejad, D. Blair, and E.L. Bibeau: Characterization of urban commuter driving profiles to optimize battery size in light-duty plug-in Electric Vehicles, *Transportation Research Part D: Transport and Environment* 16 (2011), 218-224

[6] Z. Lin, J. Ding, C. Liuand D.L. Greene: PHEV Energy Use Estimation: Validating the Gamma Distribution for Representing the Random Daily Driving Distance, in: *TRB 91st annual meeting* (2012), 1–14, Transportation Research Board `http://amonline.trb.org/1sir3o/1`.

[7] A. Blum: *Electro-mobility: statistical analysis of human mobility patterns* Master Thesis, infernum, Wuppertal, 2014.

[8] M.A. Tamor, C. Gearhart, C. Soto: A statistical approach to estimating acceptance of electric vehicles and electrification of personal transportation, *Transportation Research Part C: Emerging Technologies* 26 (2013), 125-134, `http://dx.doi.org/10.1016/j.trc.2012.07.007`.

[9] D. J. Finney: On the distribution of a variate whose logarithm is normally distributed, *J. Roy. Statist. Soc. Ser. B* 7 (1941), 155-161

[10] W.-H. Shen: Estimation of parameters of a lognormal distribution, *Taiwanese Journal of Mathematics* Vol. 2 (1998), 243-250.

[11] K.K. Hu: Notes on the Use of Propagation of Error Formulas, *Journal of reasearch of the national Bureau of Standards – C. Engineering and Instrumentation* 40C (1966), pp. 263-273.

[12] P. Plötz, T. Gnann and M. Wietschel: Total Ownership Cost Projection for the German Electric Vehicle Market with Implications for its Future Power and Electricity Demand, in *Proceedings of the 2012 Enerday Conference*, Dresden, 2012.

Matlab Codes

```
% This program estimates the probability of rare long distance
% trips by assuming a log-normal distribution of individual
% daily vehicle kilometres travelled (DVKT) and estimates
% the parameters of the log-normal distribution from a finite
% sample of the vehicle's driving. This program is mainly
% intended to test the validity and goodness of the estimates.

%% Parameters
clear all;
% The "real" parameters mu and sigma:
mu = 3.43; sig = 1.20;
% Parameters of observation
obsdays = 14; % integer multiples of 7 required: 7, 14, 21, etc
range = 150; % example electric driving range in km
daysreal = 5/7*365*(1-logncdf(range,mu,sig)); % exact value

% number of Monte Carlo/bootstrap simulation runs
bstr = 1000;

%% BOOTSTRAPPING DRIVERS
mus = zeros(bstr,1);
sigs = zeros(bstr,1);
drivdays = binornd(obsdays,5/7,bstr,1);
for ll = 1:bstr
    % suppose we draw data from the distribution
    data=lognrnd(mu,sig,drivdays(ll),1);
    % estimate parameters:
    muhat = mean(log(data));
    sighat= std(log(data));
  %save values
    mus(ll) = muhat;
    sigs(ll)=sighat;
end

% Collect estimated number of days requiring adaptation
days = drivdays/obsdays*365.*(1-logncdf(range,mus,sigs));

%% plotting
figure(1); clf;
numbins = 20;
subplot(2,2,1);
  [n, x] = hist(mus,numbins); n = n/sum(n);
  bar(x,n,0.8,'c');
subplot(2,2,2);
  [n, x] = hist(sigs,numbins); n = n/sum(n);
  bar(x,n,0.8,'c');
subplot(2,2,3);
  [n,x] = hist(days,numbins); n = n/sum(n);
  bar(x,n,0.8,'c');
subplot(2,2,4);
  plot(mus,drivdays,'+');
  xlabel('\mu');ylabel('driving days');
  [rho, pval] = corr(mus,drivdays);
  title(['p-value for corr. <> 0: '...
    ,num2str(pval*100,'%2.0f'),'%'])

disp('estimate number of days with 95% confidence interval:')
disp(['D = ', num2str(mean(days),'%3.1f'),' \pm ',...
  num2str(1.96*std(days)/sqrt(bstr),'%3.2f'),...
```

```
  ' vs. D = ',num2str(daysreal,'%3.1f')])

%% Confidence bands
% Plot number of days (estimated & 'real') vs. electric range
ranges = 50:5:400;
% need matrices for simple syntax:
daysreal = 6/7*365*(1-logncdf(ranges,mu,sig));
mumat = mus*ones(1,length(ranges));
sigmat = sigs*ones(1,length(ranges));
daysmat = drivdays*ones(1,length(ranges));
rangemat = ones(length(mus),1)*ranges;

% compute matrix of days requiring adaption:
daysmat = daysmat/obsdays.*365.*(1-logncdf(rangemat,mumat,sigmat));

% collect statistics:
daystat = quantile(daysmat,[0.25 0.5 0.75],1);
daysmean = mean(daysmat,1);

% plot estimated number of long distance trips and 'real' values
figure(2); clf;
    plot(ranges,daystat(1,:),'--',ranges,daystat(2,:),'b-',ranges,daystat(3,:),'b--'); hold on;
    plot(ranges,daysmean,'k--',ranges,daysreal,'r-','LineWidth',2); hold off;
legend('1st quartile','median','3rd quartile','mean','real value')
axis([0 400 0 80]); grid on;
xlabel('electric range [km]')
ylabel('days in year requiring adaption')

%% BOOTSTRAPPING OBSERVATION PERIOD
weeks = 1:52;
mus = zeros(bstr,max(weeks));
sigs = zeros(bstr,max(weeks));
drivdays = zeros(bstr,max(weeks));
obsdays = ones(bstr,max(weeks));
for w = weeks
    obsdays(:,w) = 7*w;
    drivdays(:,w) = binornd(obsdays(:,w),5/7,bstr,1);
    for ll = 1:bstr
        % suppose we draw data from the distribution
        data=lognrnd(mu,sig,drivdays(ll,w),1);
        % estimate parameters:
        muhat = mean(log(data));
        sighat= std(log(data));
      %save values
        mus(ll,w) = muhat;
        sigs(ll,w)=sighat;
    end
end

days = drivdays./obsdays*365.*(1-logncdf(range,mus,sigs));
days1 = mean(days,1); days2 = mean(days,2);
quant = quantile(days,[1/6 0.5 5/6],1);
iqr = quant(3,:)-quant(1,:);

%% plotting the error sigma_D
figure(3); clf;
subplot(2,4,1); hist(days(:,1));
subplot(2,4,2); hist(days(:,4));
subplot(2,4,5); hist(days(:,12));
subplot(2,4,6); hist(days(:,52));
```

```
subplot(2,4,[3 4 7 8]);
  F = logncdf(range,mu,sig);
  f = lognpdf(range,mu,sig);
  p = 5/7;
  % compute estimated error
  factor = (1-F)^2*(1-p)+...
      (range*f*sig).^2*(1+((log(range)-mu)./(sqrt(2)*sig)).^2);
  theory = 2*365*sqrt(p./(7*weeks))*sqrt(factor);
  plot(7*weeks,iqr,'o',7*weeks,theory,'r--');
  xlabel('observation period N [days]');
  ylabel(' 2 \sigma_D [days]');

%% plotting the mean with confidence band
figure(4); clf;
  est = quant(2,:);
  up= est+iqr/2;
  low=est-iqr/2;
  errorbar(7*weeks,est,est-low,est-up,'o'); hold on;
  D0 = 5/7*365*(1-logncdf(range,mu,sig));
  plot(7*weeks,D0*ones(1,max(weeks)),'r-');
  hold off;
  xlabel('observation period N [days]');
  ylabel(' D \pm \sigma_D [days]');
  axis([0 380 0 50]);
  legend('simulation','real value')

%% plotting the error in VKT sigma_R/R
% annual VKT
Rnum = mean(365*p*exp(mus+sigs.^2/2),1);
Rhat = 365*p*exp(mu+sig^2/2);
% Standard error annual VKT
qR = quantile(365*p*exp(mus+sigs.^2/2),[normcdf(-1,0,1) 0.5 normcdf(+1,0,1)],1);
sigmaRnum = (qR(3,:)-qR(1,:))/2;
sigmaRhat = Rhat./sqrt(p*7*weeks)*sqrt(1-p+sig^2+sig^4/2);

figure(5); clf;
subplot(1,3,1)
    errorbar(7*weeks,Rnum/1000,sigmaRnum/1000,sigmaRnum/1000,'o'); hold on;
    plot(7*weeks,Rhat/1000*ones(1,length(weeks)),'r--','LineWidth',2); hold off;
    xlabel('observation period N [days]');
    ylabel('Annual VKT R \pm \sigma_R [1,000 km]');


subplot(1,3,2)
plot(7*weeks,sigmaRnum,'o',7*weeks,sigmaRhat,'r--','LineWidth',2);
    xlabel('observation period N [days]');
    ylabel('Standard error \sigma_R [km]');
    axis([-10 380 0 13900]);
    set(gca,'Ytick',[0 2 4 6 8 10 12]*1e3);
    set(gca,'YTicklabel',{'0','2,000','4,000','6,000','8,000','10,000','12,000'})
    legend('simulation','theory')

subplot(1,3,3)
loglog(7*weeks,sigmaRnum./Rnum,'o',...
  7*weeks,sigmaRhat./Rhat,'r--','LineWidth',2);
    xlabel('observation period N [days]');
    ylabel('realtive error \sigma_R / R ');
    axis([5 550 0.07 1]); grid on
```

Authors' affiliations

Patrick Plötz

Fraunhofer Institute for Systems and Innovation Research (Fraunhofer ISI)
Competence Center Energy Technology and Energy Systems

Contact: Brigitte Kallfass

Fraunhofer Institute for Systems
and Innovation Research (Fraunhofer ISI)
Breslauer Strasse 48
76139 Karlsruhe
Germany
Phone: +49 / 721 / 6809-150
Fax: +49 / 721 / 6809-272
E-Mail: brigitte.kallfass@isi.fraunhofer.de
www.isi.fraunhofer.de

Karlsruhe 2014