Intermittent state-space model for demand forecasting

Ivan Svetunkov and John Boylan

19th IIF Workshop

29th June 2016

Lancaster University Management School | Lancaster Centre for Forecasting

LCF

Motivation

Croston (1972) proposes a method for intermittent demand forecasting, mentioning the model: $y_\iota = x_t \cdot z_t$

He estimates probability using intervals between demands ($\frac{1}{q_t}$).

He also assumes that probability is constant between occurrences.

Syntetos and Boylan (2001, 2005) show that the conditional expectation of Croston's method is biased.

They propose an approximation, that corrects the error.

LCF

Motivation

Snyder (2002) looks at Croston's method in details, claiming that the underlying model is: $y_\iota = x_t \cdot z_t + \epsilon_t$.

This model produces both positive and negative data.

This is a drawback, so Snyder (2002) proposes a modification, taking $\exp$ of non-zero demands.

Motivation

Shenstone and Hyndman (2005) study several additive models, possibly underlying Croston's method.

They argue that any model underlying Croston's method must be:

- non-stationary,
- defined on continuous space.

They conclude that the implied model has non-realistic properties.

They support Snyder (2002) approach with $\exp$.

Motivation

Teunter et al. (2011) propose a model taking inventory obsolescence into account.

The probability of having a demand is decreasing when demand does not occur.

Simulation is done, but estimation of parameters is skipped.

In the following paper Zied Babai et al. (2014) optimise several methods, including TSB.

They use MSE calculated as a difference between the estimate and the actual demand.

Motivation

Kourentzes (2014) investigates the estimation of Croston, SBA, TSB.

He discusses several cost functions.

And proposes two new ones, which improves estimation of methods.

He finds that optimisation of initial states increases forecasting accuracy.

Motivation, overall

There is no concise model, underlying all the methods.

Because of Shenstone and Hyndman (2005) we believe that it doesn't exist.

Intermittent demand methods are disconnected from slow-moving data methods.

And we still need to make good decisions about replenishment levels.

Universal model

Universal model

Very general model:

$$y_t = o_t \widetilde{y}_t, \tag{1}$$

where $o_t \sim \text{Bernoulli}(p_t)$ and $\widetilde{y}_t$ is a statistical model of our choice.

This corresponds to Croston's original idea.

If $o_t = 1,$ for any t, then this is slow-moving data model.

Additive state-space model (Snyder, 1985)

State-space model:

$$\begin{matrix} y_t = o_t(w' v_{t-1} + \epsilon_t) \\ v_t = F v_{t-1} + g \epsilon_t \end{matrix}, \quad (2)$$

v_{t-1} vector of states, w is measurement vector, F is transition matrix, g is persistence vector, $\epsilon_t \sim \mathrm{N}(0, \sigma^2)$.

Example. iETS(A,N,N) with constant probability:

$$\begin{matrix} y_t = o_t(l_{t-1} + \epsilon_t) \\ l_t = l_{t-1} + \alpha \epsilon_t \end{matrix}, \quad (3)$$

where $o_t \sim \text{Bernoulli}(p)$.

General state-space (based on Hyndman et al. (2008))

State-space model for any ETS:

$$\begin{aligned} y_t &= o_t \left(w(v_{t-1}) + r(v_{t-1}, \epsilon_t) \right) \\ v_t &= F(v_{t-1}) + g(v_{t-1}, \epsilon_t) \end{aligned} . \quad (4)$$

Example. iETS(M,Ad,N) with constant probability:

$$\begin{aligned} y_t &= o_t(l_{t-1} + \phi b_{t-1})(1 + \epsilon_t) \\ l_t &= (l_{t-1} + \phi b_{t-1})(1 + \alpha \epsilon_t) \\ b_t &= \phi b_{t-1}(1 + \beta \epsilon_t) \end{aligned} , \quad (5)$$

where $o_t \sim \text{Bernoulli}(p)$, $(1 + \epsilon_t) \sim \log \text{N}(0, \sigma^2)$.

Advantages

What are the advantages of such a model?

- Statistical rationale for intermittent demand;
- Connection between conventional and intermittent models;
- Correct estimation of mean;
- Simpler variance estimation;
- Prediction intervals;

Advantages

What else?

- Both additive and multiplicative ETS models;
- Any statistical model;
- Likelihood function;
- Solution to initialisation and optimisation problems;
- Model selection.

Disadvantages

What are the disadvantages of such a model?

- May need more observations...
- ...Especially for trend and seasonal models;
- Derivations in some cases may be messy.

iETS(M,N,N),

constant probability

iETS(M,N,N), constant probability

iETS(M,N,N) model has the form:

$$\begin{array}{l} y_t = o_t l_{t-1}(1+\epsilon_t) \\ l_t = l_{t-1}(1+\alpha\epsilon_t) \end{array}, \quad (6)$$

where $o_t \sim \text{Bernoulli}(p)$.

iETS(M,N,N) underlies SES (Hyndman et al., 2008).

Conditional expectation:

$$E(y_{t+h}|t) = pE(\tilde{y}_{t+h}|t) = pw'F^{h-1}v_t = pl_t.$$

LCF

iETS(M,N,N), constant probability

Conditional variance:

$$V(y_{t+h}|t) = p(1-p)l_t^2 + pl_t^2\sigma^2 \left(1 + \alpha^2(1+\sigma^2)\sum_{j=1}^{h-1}(1+\alpha^2\sigma^2)\right).$$

Messy because of the multiplicative error.

iETS(M,N,N), constant probability

Likelihood can be derived taking probabilities:

$P(y_t|o_t = 1, \theta, \sigma^2) = p\frac{1}{y_t}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(1+\epsilon_t)^2}{2\sigma^2}}$,
$P(y_t|o_t = 0, \theta, \sigma^2) = 1 - p$.

Product of all the zero and non-zero cases is then:

$$L(\theta, \sigma^2|y_t) = \prod_{o_t=1} p\frac{1}{y_t}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(1+\epsilon_t)^2}{2\sigma^2}} \prod_{o_t=0}(1-p). \qquad (7)$$

LCF

iETS(M,N,N), constant probability

The concentrated log-likelihood is simple:

$$\ell(\theta, \hat{\sigma}^2 | y_t) = -\frac{T_1}{2} \left(\log\left(2\pi e\right) + \log\left(\hat{\sigma}^2\right)\right) - \sum_{o_t=1} \log(y_t) \\ + T_0 \log(1-p) + T_1 \log p, \tag{8}$$

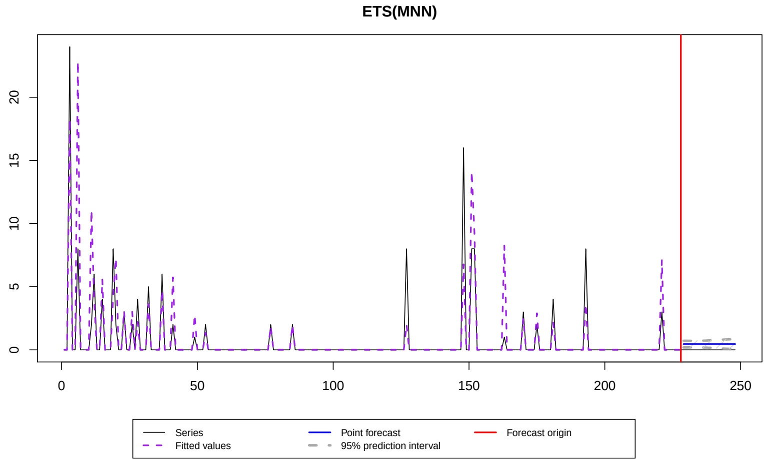
where T is number of all observations, T_0 is number of zeroes, T_1 number of non-zero demands.

The variance of the error estimated using likelihood (8) is:

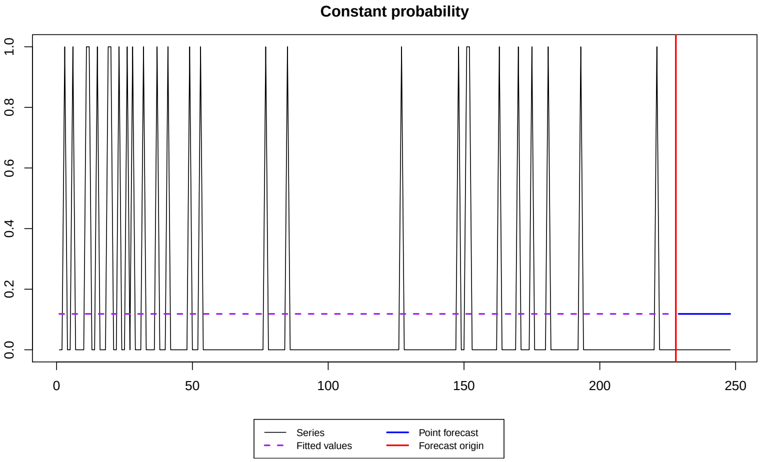
$$\hat{\sigma}^2 = \frac{1}{T_1} \sum_{o_t=1} (1+\epsilon_t).$$

The probability can also be derived from (8): $p = \frac{T_1}{T}$.

Example. Intermittent demand



Example. Probabilities



Simple iETS. Sub-conclusion

- Pretty easy statistical model;
- Multiplicative ETS is possible and makes more sense than additive;
- But probability is currently constant;

LCF

iETS(M,N,N),

time varying probability,

Croston style

Croston's iETS(M,N,N)

ETS(M,N,N) + compound Bernoulli distribution:
$o_t \sim \text{Bernoulli}(p_t)$, where $p_t = \frac{1}{1+q_t}$,
q_t are intervals between demands. If $q_t = 0$, then $p_t = 1$.

Assumption: Probability changes only when demand occurs.

State-space model for probabilities:

$$\begin{aligned} q_t &= l_{q,t-1}(1+\varepsilon_t) \\ l_{q,t} &= l_{q,t-1}(1+\delta\varepsilon_t) \end{aligned}, \tag{9}$$

where $(1+\varepsilon_t) \sim \log \text{N}(0, \sigma_q^2)$

Croston's iETS(M,N,N)

Overall iETS(M,N,N) Croston style is:

$$
\begin{aligned}
& y_t = o_t l_{t-1}(1+\epsilon_t) \\
& l_t = l_{t-1}(1+\alpha\epsilon_t) \\
& \\
& q_t = l_{q,t-1}(1+\varepsilon_t) \\
& l_{q,t} = l_{q,t-1}(1+\delta\varepsilon_t) \\
& \\
& (1+\epsilon_t) \sim \log \mathrm{N}(0,\sigma^2) \\
& o_t \sim \text{Bernoulli}(\tfrac{1}{1+q_t}) \\
& (1+\varepsilon_t) \sim \log \mathrm{N}(0,\sigma_q^2).
\end{aligned} \tag{10}
$$

Now it becomes a bit more complicated...

Croston's iETS(M,N,N)

Conditional expectation:

$$E(y_{t+h}|t) = l_t E\left(\frac{1}{1+q_{t+h}}\middle| t\right).$$

Not yet simplified:

$$E(y_{t+h}|t) = l_t E\left(\frac{1}{1+l_{q,t}\prod_{j=1}^{h-1}(1+\delta\varepsilon_{t+j})(1+\varepsilon_{t+h})}\middle| t\right).$$

We feel that this should be close to SBA.

Croston's iETS(M,N,N)

Variance is currently mind blowing...

But it should be based on the variance of o_t:
$\sigma_o^2 = p_t(1 - p_t)$

Meaning that the conditional variance of y_{t+h} is:

$$V(y_{t+h}|t) = E\left(\frac{1}{1+q_{t+h}}\middle| t\right)\left(1 - E\left(\frac{1}{1+q_{t+h}}\middle| t\right)\right) l_t^2 + E\left(\frac{1}{1+q_{t+h}}\middle| t\right) l_t^2 \sigma^2 \left(1 + \alpha^2 (1+\sigma^2) \sum_{j=1}^{h-1} (1+\alpha^2\sigma^2)\right).$$

LCF

Croston's iETS(M,N,N)

Likelihood however can be done in two stages (assuming demand sizes and intervals are independent):

1. Likelihood for intervals;
2. Likelihood for demands.

Both of them are based on lognormal distributions.

Croston's iETS(M,N,N)

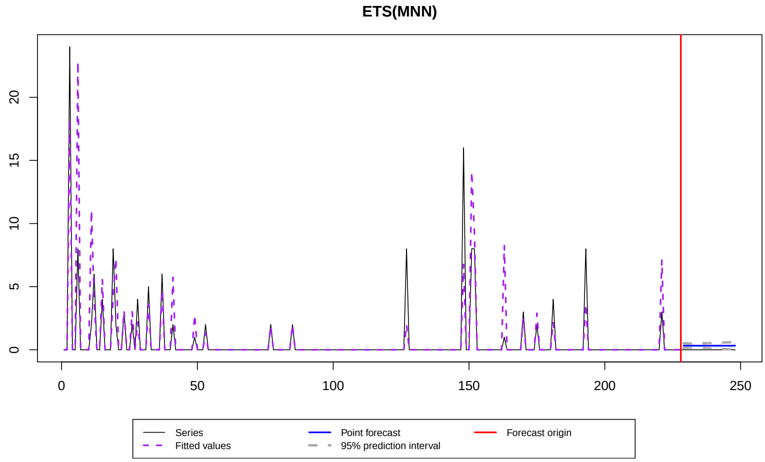
Concentrated log-likelihoods.
For intervals (first stage):

$$\ell(\theta_q, \hat{\sigma}_q^2 | q_t) = -\frac{T_q}{2} \left(\log \left(2\pi e \right) + \log \left(\hat{\sigma}_q^2 \right) \right) - \sum_{t=1}^{T_q} \log(q_t), \quad (11)$$

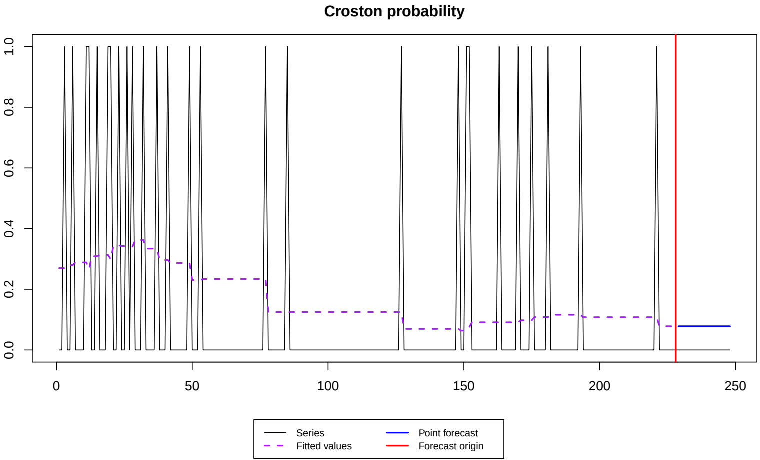
For demands (second stage):

$$\begin{aligned} \ell(\theta, \hat{\sigma}^2 | y_t) = & -\tfrac{T_1}{2} \left(\log \left(2\pi e \right) + \log \left(\hat{\sigma}^2 \right) \right) - \sum_{o_t=1} \log(y_t) \\ & + \sum_{o_t=0} \log(1 - p_t) + \sum_{o_t=1} \log p_t , \end{aligned} \quad (12)$$

Croston's iETS(M,N,N). Example



Croston's iETS(M,N,N). Example. Probabilities



Croston's iETS. Sub-conclusion

- There is a statistical model underlying Croston's method;
- Conditional expectation should be closer to SBA;
- Conditional variance can be found analytically;
- Probabilities are updated only when demand occurs;
- There are still some problems with derivations.

iETS(M,N,N),

time varying probability,

TSB

TSB iETS(M,N,N)

ETS(M,N,N) + compound Bernoulli distribution:
$o_t \sim$ Bernoulli(p_t), where:

$$\begin{array}{l} p_t = l_{p,t-1}(1+\xi_t) \\ l_{p,t} = l_{p,t-1}(1+\delta\xi_t) \end{array}. \quad (13)$$

p_t can be estimated as naïve probability: $p_t = o_t$.

We want to have conditional Beta(a,b) distribution.

But this means that $p_t \in (0,1)$.

We need boundary values!

TSB iETS(M,N,N)

Temporary fix – simple transfer function:

$$p'_t = (1 - 2\kappa)p_t + \kappa,$$

where κ is some small number. e.g. $\kappa = 10^{-20}$.

This means that $p'_t \in (\kappa, 1 - \kappa)$.

So $p'_t \sim$ Beta(a,b).

TSB iETS(M,N,N)

The fixed TSB iETS(M,N,N) is then:

$$
\begin{aligned}
& y_t = o_t l_{t-1}(1+\epsilon_t) \\
& l_t = l_{t-1}(1+\alpha\epsilon_t) \\
& \\
& p_t = \frac{p'_t - \kappa}{1-2\kappa} \\
& p'_t = l_{p,t-1}(1+\xi_t) \\
& l_{p,t} = l_{p,t-1}(1+\delta\xi_t) \\
& \\
& (1+\epsilon_t) \sim \log \mathrm{N}(0,\sigma^2) \\
& o_t \sim \text{Bernoulli}(p_t) \\
& p'_t \sim \text{Beta(a,b)}
\end{aligned} . \qquad (14)
$$

TSB iETS(M,N,N)

Conditional expectation is simpler than in Croston:

$$E(y_{t+h}|t) = l_t \frac{l_{p,t-1} - \kappa}{1 - 2\kappa}.$$

Conditional variance is based on Bernoulli $p_{t+h|t}(1 - p_{t+h|t})$:

$$\begin{aligned} V(y_{t+h}|t) = & \tfrac{l_{p,t-1}-\kappa}{1-2\kappa} \left(1 - \tfrac{l_{p,t-1}-\kappa}{1-2\kappa}\right) l_t^2 \\ & + \tfrac{l_{p,t-1}-\kappa}{1-2\kappa} l_t^2 \sigma^2 \left(1 + \alpha^2(1+\sigma^2) \sum_{j=1}^{h-1} (1 + \alpha^2\sigma^2)\right). \end{aligned}$$

TSB iETS(M,N,N)

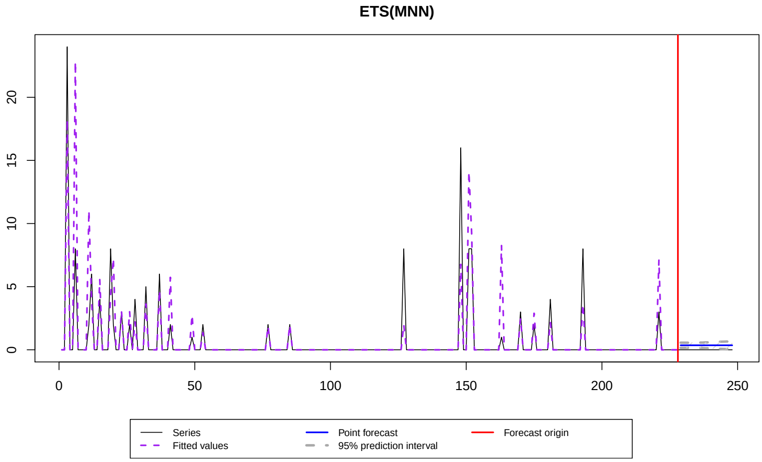
Concentrated log-likelihood in two stages.
For the probability (stage 1):

$$\begin{aligned}\ell(\theta_p, a, b | p_t) = & (a-1)\sum_{t=1}^{T} \log(l_{p,t-1}(1+\xi_t)) \\ & +(b-1)\sum_{t=1}^{T} \log(1 - l_{p,t-1}(1+\xi_t)) \\ & -T \log B(a,b),\end{aligned} \tag{15}$$

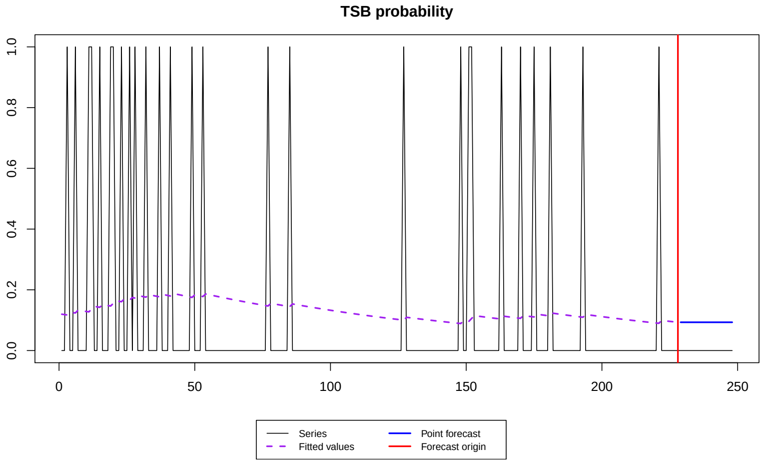
For the demand sizes (stage 2):

$$\begin{aligned}\ell(\theta, \hat{\sigma}^2 | y_t) = & -\frac{T_1}{2}\left(\log(2\pi e) + \log\left(\hat{\sigma}^2\right)\right) - \sum_{o_t=1} \log(y_t) \\ & + \sum_{o_t=0} \log(1-p_t) + \sum_{o_t=1} \log p_t,\end{aligned} \tag{16}$$

TSB iETS(M,N,N). Example



TSB iETS(M,N,N). Example. Probabilities



TSB iETS. Sub-conclusion

- There is a statistical model underlying TSB;
- Estimation problem solved;
- Works fine even with the proposed approximation;
- p_t is unknown, problem with estimation;
- Problem with distribution of p_t;
- Multiplicative damped trend could be more appropriate.

Real time series example

Example on the real data

1. 58 intermittent time series,
2. One product, different branches, daily data,
3. 248 observations each, 10 – 103 demand occurrences,
4. Holdout sample of 20 obs,
5. iETS using "es" from "smooth" package in R (`https://github.com/config-i1/smooth`):
    - Stable probability,
    - Croston's probability,
    - TSB probability,
6. Croston's method and TSB, "tsintermittent" package in R.

Example on the real data

Method	sPIS	sAPIS	ARMSE	Complex bias
iETS, stable	-609.2	2219.6	1.00	-46.3%
iETS, Croston	-442.0	2299.4	0.99	-48.4%
iETS, TSB	-538.2	**2082.3**	**0.92**	**-46.1%**
Croston's method	**-256.0**	2158.9	1.03	-53.2%
TSB method	-279.6	2116.2	1.03	-52.8%
Zero forecast	-2363.6	2363.6	0.82	99.5%

Table: Intermittent demand data performance.

Conclusions

Conclusions

- We proposed a very simple modification, that can be applied to any model;
- iETS is one of such models;
- Multiplicative models are available now;
- Model selection is also available;
- It can even be done between Stable / Croston / TSB;

Conclusions

- Conditional expectation can be correctly estimated;
- The same holds for the conditional variance;
- Prediction intervals for intermittent data;
- Croston and TSB have underlying iETS model;
- Estimation problem is now solved for them.

Thank you for your attention!

Ivan Svetunkov, John Boylan

i.svetunkov@lancaster.ac.uk
j.boylan@lancaster.ac.uk

Lancaster University Management School
Lancaster Centre for Forecasting

Croston, J. D., 1972. Forecasting and Stock Control for Intermittent Demands. Operational Research Quarterly (1970-1977) 23 (3), 289.

Hyndman, R. J., Koehler, A., Ord, K., Snyder, R., 2008. Forecasting with Exponential Smoothing. Springer Berlin Heidelberg.

Kourentzes, N., 2014. On intermittent demand model optimisation and selection. International Journal of Production Economics 156, 180–190.

Shenstone, L., Hyndman, R. J., 2005. Stochastic models underlying Croston's method for intermittent demand forecasting. Journal of Forecasting 24 (6), 389–402.

Snyder, R., 2002. Forecasting sales of slow and fast moving inventories. European Journal of Operational Research 140 (3), 684–699.

Snyder, R. D., 1985. Recursive Estimation of Dynamic Linear

Models. Journal of the Royal Statistical Society, Series B (Methodological) 47 (2), 272–276.

Syntetos, A., Boylan, J., 2001. On the bias of intermittent demand estimates. International Journal of Production Economics 71 (1-3), 457–466.

Syntetos, A. A., Boylan, J. E., 2005. The accuracy of intermittent demand estimates. International Journal of Forecasting 21 (2), 303–314.

Teunter, R. H., Syntetos, A. A., Zied Babai, M., 2011. Intermittent demand: Linking forecasting to inventory obsolescence. European Journal of Operational Research 214 (3), 606–615.

Zied Babai, M., Syntetos, A., Teunter, R., 2014. Intermittent demand forecasting: An empirical study on accuracy and the risk of obsolescence. International Journal of Production Economics 157 (1), 212–219.