Non-life insurance mathematics

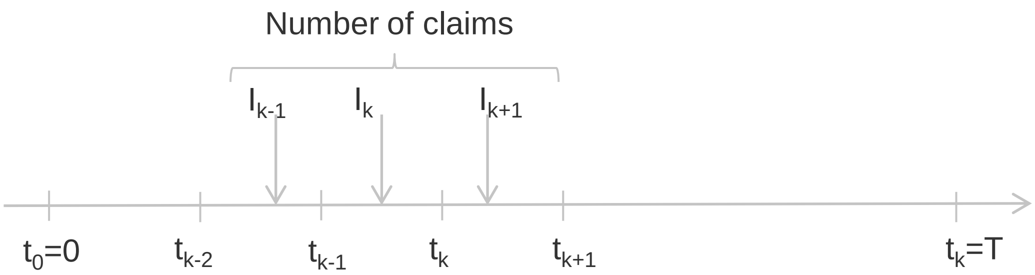
Nils F. Haavardsson, University of Oslo and DNB Skadeforsikring

Insurance mathematics is fundamental in insurance economics

The result drivers of insurance economics:

Result elements:	**Result drivers:**
+ Insurance premium	Risk based pricing, reinsurance
+ financial income	International economy for example interest rate level, risk profile for example stocks/no stocks
- claims	risk reducing measures (for example installing burglar alarm), risk selection (client behaviour), change in legislation, weather phenomenons, demographic factors, reinsurance
- operational costs	measures to increase operational efficiency, IT-systems, wage development
= result to be distributed among the owners and the	Tax politics

The world of Poisson (Chapter 8.2)



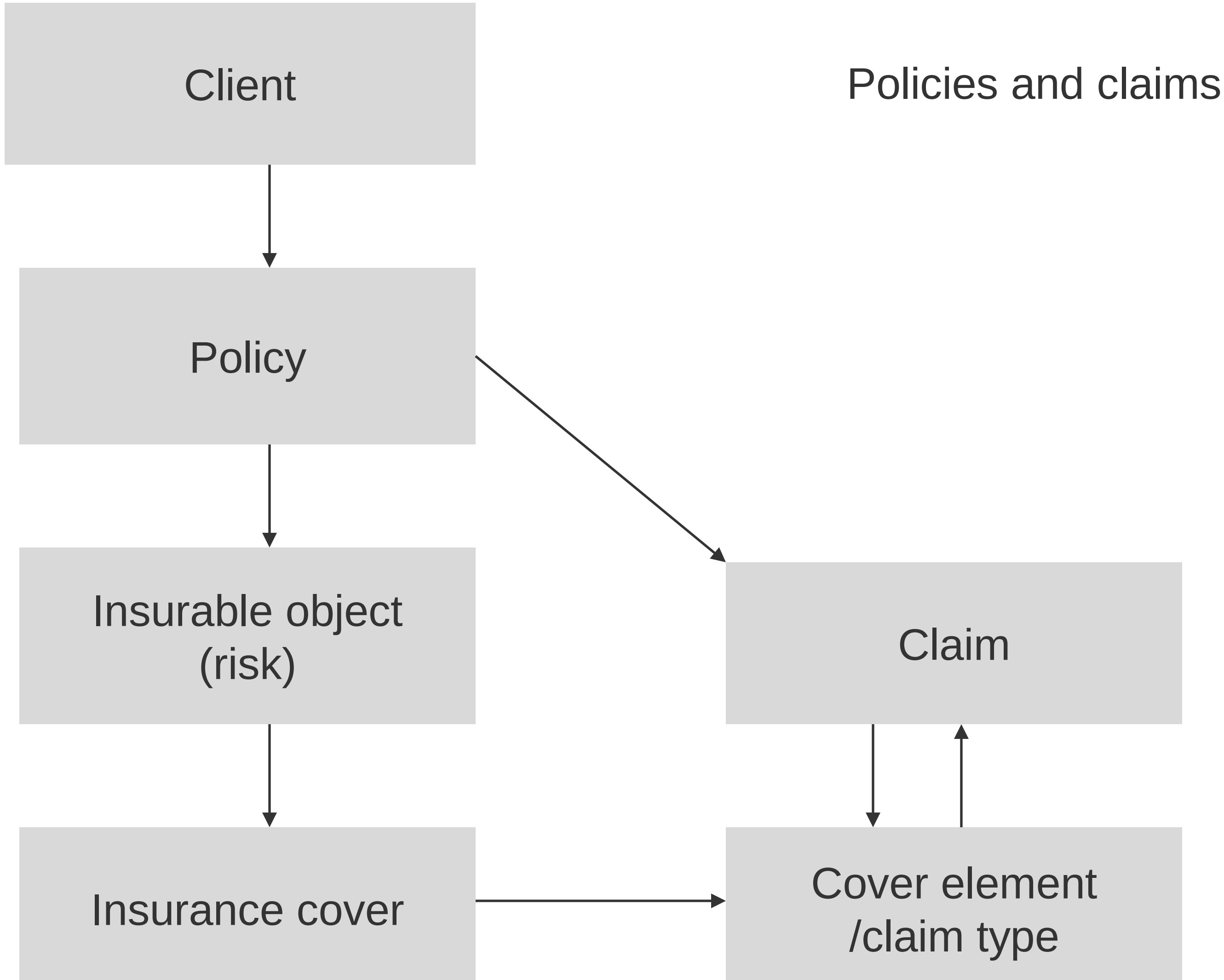
- What is rare can be described mathematically by cutting a given time period T into K small pieces of equal length h=T/K
- On short intervals the chance of more than one incident is remote
- Assuming no more than 1 event per interval the count for the entire period is

$N=I_1+\ldots+I_K$, where I_j is either 0 or 1 for j=1,...,K

- If $p=\Pr(I_k=1)$ is equal for all k and events are independent, this is an ordinary Bernoulli series

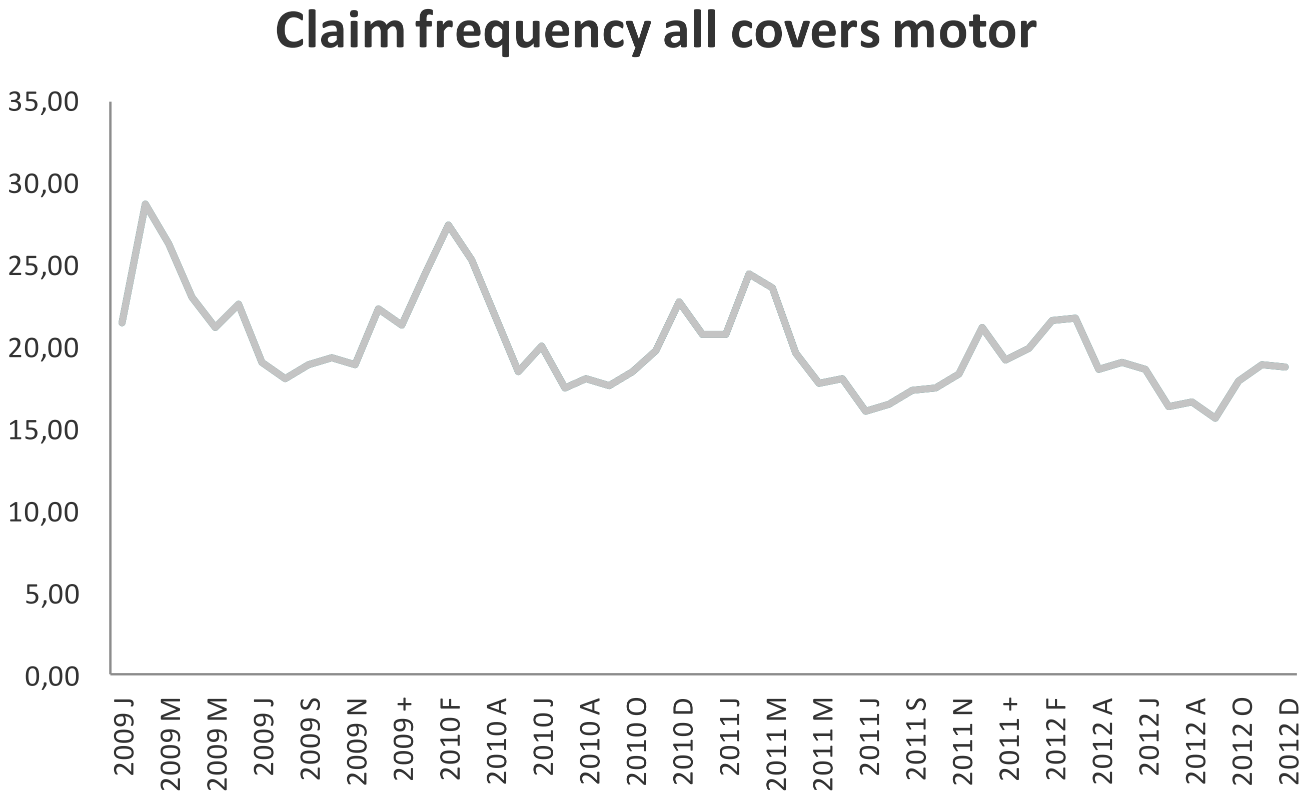
$$\Pr(N=n)=\frac{K!}{n!(K-n)!}p^n(1-p)^{K-n}, \quad \text{for} \quad n=0,1,\ldots,K$$

- Assume that p is proportional to h and set $p=\mu h$ where μ is an intensity which applies per time unit


Policies and claims
Client
Policy
Insurable object (risk)
Insurance cover
Claim
Cover element /claim type

Key ratios – claim frequency

- The graph shows claim frequency for all covers for motor insurance
- Notice seasonal variations, due to changing weather condition throughout the years



Random intensities (Chapter 8.3)

- How μ varies over the portfolio can partially be described by observables such as age or sex of the individual (treated in Chapter 8.4)
- There are however factors that have impact on the risk which the company can't know much about
  - Driver ability, personal risk averseness,
- This randomeness can be managed by making μ a stochastic variable
- This extension may serve to capture uncertainty affecting all policy holders jointly, as well, such as altering weather conditions
- The models are conditional ones of the form

$$N \mid \mu \sim Poisson(\mu T) \quad \text{and} \quad \mathbf{N} \mid \mu \sim Poisson(J\mu T)$$

Policy level — Portfolio level

- Let $\xi = E(\mu)$ and $\sigma = \text{sd}(\mu)$ and recall that $E(N \mid \mu) = \text{var}(N \mid \mu) = \mu T$ which by double rules in Section 6.3 imply

$$E(N) = E(\mu T) = \xi T \quad \text{and} \quad \text{var}(N) = E(\mu T) + \text{var}(\mu T) = \xi T + \sigma^2 T^2$$

- Now E(N)<var(N) and N is no longer Poisson distributed

Overview

Important issues	Models treated	Curriculum	Duration (in lectures)
What is driving the result of a non-life insurance company?	insurance economics models	Lecture notes	0,5
How is claim frequency modelled?	Poisson, Compound Poisson and Poisson regression	Section 8.2-4 EB	1,5
How can claims reserving be modelled?	Chain ladder, Bernhuetter Ferguson, Cape Cod,	Note by Patrick Dahl	2
How can claim size be modelled?	Gamma distribution, log-normal distribution	Chapter 9 EB	2
How are insurance policies priced?	Generalized Linear models, estimation, testing and modelling. CRM models.	Chapter 10 EB	2
Credibility theory	Buhlmann Straub	Chapter 10 EB	1
Reinsurance		Chapter 10 EB	1
Solvency		Chapter 10 EB	1
Repetition			1

Overview of this session

- What is a fair price of insurance policy?
- The model (Section 8.4 EB)
- An example
- Why is a regression model needed?
- Repetition of important concepts in GLM

What is a fair price of an insurance policy?

- Before ”Fairness” was supervised by the authorities (Finanstilsynet)
  - To some extent common tariffs between companies
  - The market was controlled
- During 1990’s: deregulation
- Now: free market competition supposed to give fairness
- According to economic theory there is no profit in a free market (in Norway general insurance is cyclical)
- Hence, the price equals the expected cost for insurer
- Note: cost of capital may be included here, but no additional profit
- Ethical dilemma:
  - Original insurance idea: One price for all
  - Today: the development is heading towards micropricing
  - These two represent extremes

Expected cost

- Main component is expected loss (claim cost)
- The average loss for a large portfolio will be close to the mathematical expectation (by the law of large numbers)
- So expected loss is the basis of the price
- Varies between insurance policies
- Hence the market price will vary too
- The add other income and costs, incl administrative cost and capital cost

Adverse selection

- Too high premium for some policies results in loss of good policies to competitors
- Too low premium for some policies gives inflow of unprofitable policies
- This will force the company to charge a fair premium

Rating factors

- How to find the expected loss of every insurance policy?
- We cannot price individual policies (why?)
  - Therefore policies are grouped by rating variables
- Rating variables (age) are transformed to rating factors (age classes)
- Rating factors are in most cases categorical

The model (Section 8.4)

- The idea is to attribute variation in μ to variations in a set of observable variables $x_1,\ldots,x_v$. Poisson regressjon makes use of relationships of the form

$$\log(\mu) = b_0 + b_1 x_1 + \ldots + b_v x_v \qquad (1.12)$$

- Why $\log(\mu)$ and not μ itself?
- The expected number of claims is non-negative, where as the predictor on the right of (1.12) can be anything on the real line
- It makes more sense to transform μ so that the left and right side of (1.12) are more in line with each other.

- Historical data are of the following form

n_1	T_1	$x_{11}\ldots x_{1x}$
n_2	T_2	$x_{21}\ldots x_{2x}$
n_n	T_n	$x_{n1}\ldots x_{nv}$
Claims	exposure	covariates

- The coefficients $b_0,\ldots,b_v$ are usually determined by likelihood estimation

The model (Section 8.4)

- In likelihood estimation it is assumed that n_j is Poisson distributed $\lambda_j = \mu_j T_j$ where μ_j is tied to covariates $x_{j1},\dots,x_{jv}$ as in (1.12). The density function of n_j is then

$$f(n_j) = \frac{(\mu_j T_j)^{n_j}}{n_j!}\exp(-\mu_j T_j)$$

or

$$\log(f(n_j)) = n_j\log(\mu_j) + n_j\log(T_j) - \log(n_j!) - \mu_j T_j$$

- $\log(f(n_j))$ above is to be added over all j for the likehood function $L(b_0,\dots,b_v)$.
- Skip the middle terms n_jT_j and log $(n_j!)$ since they are constants in this context.
- Then the likelihood criterion becomes

$$L(b_0,\dots,b_v) = \sum_{j=1}^{n}\{n_j\log(\mu_j) - \mu_j T_j\} \text{ where } \log(\mu_j) = b_0 + b_1 x_{j1} + \dots + b_j x_{jv} \quad (1.13)$$

- Numerical software is used to optimize (1.13).
- McCullagh and Nelder (1989) proved that $L(b_0,\dots,b_v)$ is a convex surface with a single maximum
- Therefore optimization is straight forward.

Poisson regression: an example, bus insurance

Rating factor	class	class description
Bus age	0	0 years
	1	1-2 years
	2	3-4 years
	3	5-6 years
	4	> 6 years
District	1	central and sem-central parts of sweden's three largeest cities
	2	suburbs and middle-sized towns
	3	lesser towns, except those in 5 or 7
	4	small towns and countryside, except 5-7
	5	northern towns
	6	northern countryside
	7	gotland

- The model becomes

$$\log(\mu_j) = b_0 + b_{busage}(l) + b_{district}(s)$$

for l=1,...,5 and s=1,2,3,4,5,6,7.

- To avoid over-parameterization put $b_{\text{bus age}}(5)=b_{\text{district}}(4)=0$ (the largest group is often used as reference)

Take a look at the data first

Zone	Bus age	Duration	Number of claims	Claims Cost	Claims frequency	Claims severity	Pure premium
1	0	28	20	155 312	72,6 %	7 766	5 638
1	1	30	8	55 012	27,0 %	6 877	1 857
1	2	47	15	52 401	32,1 %	3 493	1 120
1	3	85	24	79 466	28,3 %	3 311	939
1	4	222	41	220 381	18,5 %	5 375	994
2	0	64	18	37 066	28,0 %	2 059	577
2	1	55	28	83 913	50,8 %	2 997	1 523
2	2	55	15	45 321	27,1 %	3 021	820
2	3	67	25	341 384	37,5 %	13 655	5 116
2	4	507	166	2 319 807	32,7 %	13 975	4 574
3	0	74	12	192 547	16,2 %	16 046	2 600
3	1	68	19	151 747	28,0 %	7 987	2 238
3	2	62	19	517 152	30,6 %	27 219	8 315
3	3	82	12	182 846	14,6 %	15 237	2 222
3	4	763	132	1 725 852	17,3 %	13 075	2 263
5	0	12	4	303 663	32,5 %	75 916	24 664
5	1	12	8	126 814	64,3 %	15 852	10 200
5	2	10	-		0,0 %	-	-
5	3	11	4	8 998	38,0 %	2 250	855
5	4	239	51	1 383 030	21,3 %	27 118	5 789
6	0	57	29	486 935	50,9 %	16 791	8 554
6	1	68	21	58 955	30,9 %	2 807	868
6	2	57	14	307 563	24,7 %	21 969	5 416
6	3	66	18	821 205	27,3 %	45 623	12 436
6	4	895	196	3 937 850	21,9 %	20 091	4 399
7	0	7	1	289 245	13,3 %	289 245	38 571
7	1	7	2	-	29,3 %	-	-
7	2	9	1	-	11,7 %	-	-
7	3	9	3	53 751	32,4 %	17 917	5 814
7	4	87	7	-	8,1 %	-	-
9	0	320	110	481 150	34,3 %	4 374	1 502
9	1	342	125	588 172	36,5 %	4 705	1 718
9	2	440	170	2 212 900	38,6 %	13 017	5 028
9	3	444	133	1 548 812	29,9 %	11 645	3 486
9	4	4 263	754	11 941 355	17,7 %	15 837	2 801

min pp	-
max pp	38 571
median pp	2 263
min cf	0,0 %
max cf	72,6 %
median cf	28,3 %

Then a model is fitted with some software (sas below)

Model Information	
Data Set	WORK.BUSSKUNDER
Distribution	Poisson
Link Function	Log
Dependent Variable	skadfre
Scale Weight Variable	dur

Number of Observations Read	1533
Number of Observations Used	1533
Sum of Weights	9563.658

Class Level Information		
Class	Levels	Values
zon	7	1 2 3 5 6 7 9
bussald	5	0 1 2 3 4

Criteria For Assessing Goodness Of Fit			
Criterion	DF	Value	Value/DF
Deviance	1522	2063.7560	1.3560
Scaled Deviance	1522	1415.3655	0.9299
Pearson Chi-Square	1522	2219.2407	1.4581
Scaled Pearson X2	1522	1522.0000	1.0000
Log Likelihood		-3661.7729	
Full Log Likelihood		-4768.1921	
AIC (smaller is better)		9558.3843	

Zon needs some re-grouping

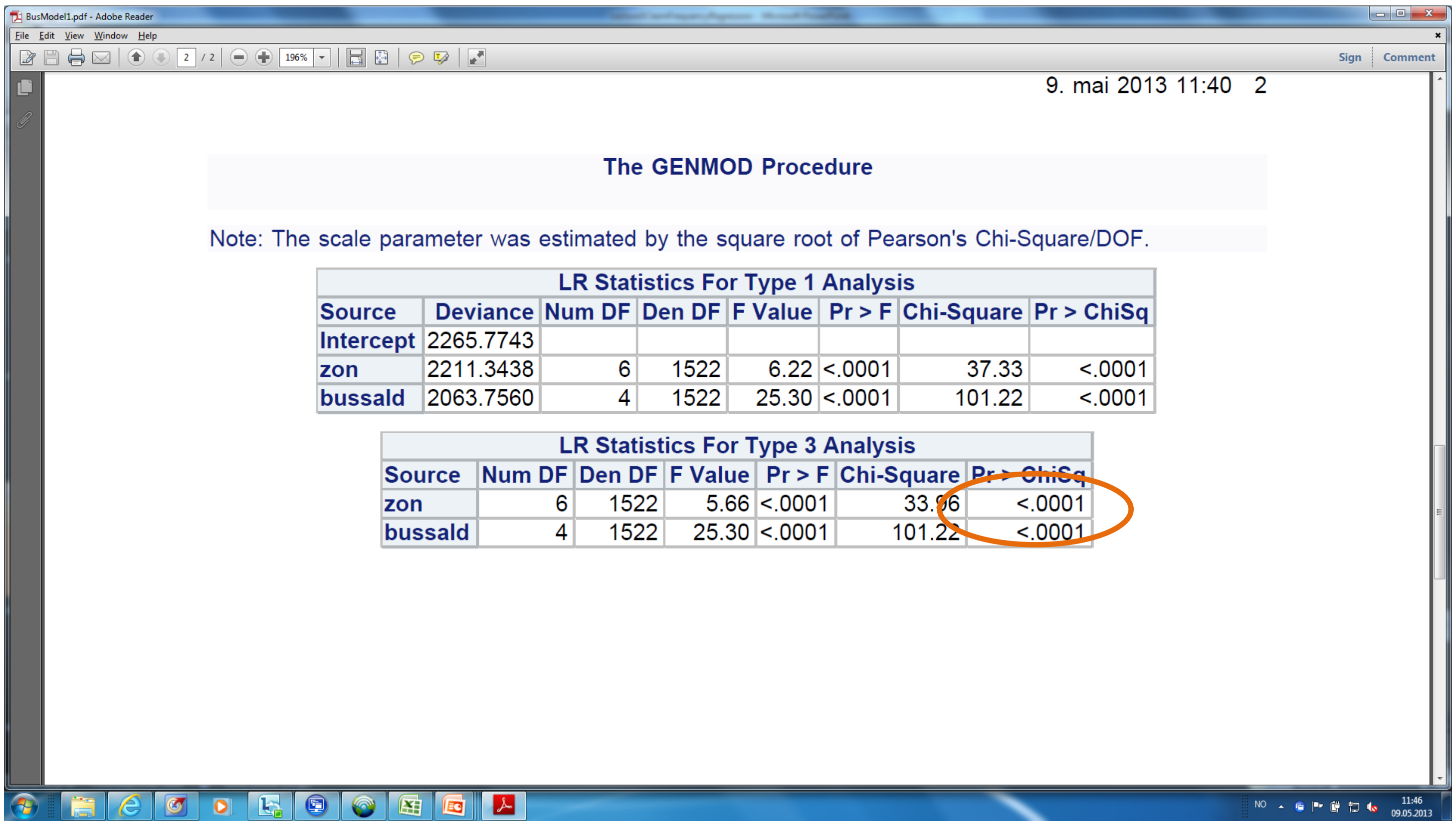
Full Log Likelihood		-4768.1921
AIC (smaller is better)		9558.3843
AICC (smaller is better)		9558.5579
BIC (smaller is better)		9617.0691

Algorithm converged.

Analysis Of Maximum Likelihood Parameter Estimates

Parameter		DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept		1	-1.6760	0.0393	-1.7530	-1.5990	1820.89	<.0001
zon	1	1	0.0782	0.1217	-0.1604	0.3168	0.41	0.5207
zon	2	1	0.3811	0.0833	0.2178	0.5443	20.94	<.0001
zon	3	1	-0.1884	0.0930	-0.3707	-0.0061	4.10	0.0428
zon	5	1	0.1191	0.1515	-0.1779	0.4160	0.62	0.4319
zon	6	1	0.1178	0.0799	-0.0389	0.2744	2.17	0.1407
zon	7	1	-0.6372	0.3245	-1.2731	-0.0012	3.86	0.0496
zon	9	0	0.0000	0.0000	0.0000	0.0000	.	.
bussald	0	1	0.5682	0.0929	0.3860	0.7503	37.38	<.0001
bussald	1	1	0.6225	0.0895	0.4470	0.7979	48.37	<.0001
bussald	2	1	0.5757	0.0858	0.4075	0.7438	45.03	<.0001
bussald	3	1	0.3896	0.0887	0.2159	0.5634	19.31	<.0001
bussald	4	0	0.0000	0.0000	0.0000	0.0000	.	.
Scale		0	1.2075	0.0000	1.2075	1.2075		

Zon and bus age are both significant



9. mai 2013 11:40 2

The GENMOD Procedure

Note: The scale parameter was estimated by the square root of Pearson's Chi-Square/DOF.

LR Statistics For Type 1 Analysis							
Source	Deviance	Num DF	Den DF	F Value	Pr > F	Chi-Square	Pr > ChiSq
Intercept	2265.7743						
zon	2211.3438	6	1522	6.22	<.0001	37.33	<.0001
bussald	2063.7560	4	1522	25.30	<.0001	101.22	<.0001

LR Statistics For Type 3 Analysis						
Source	Num DF	Den DF	F Value	Pr > F	Chi-Square	Pr > ChiSq
zon	6	1522	5.66	<.0001	33.96	<.0001
bussald	4	1522	25.30	<.0001	101.22	<.0001

Model and actual frequencies are compared

parameter	level 1	estimate
Intercept		0,19
zon	1	1,08
zon	2	1,46
zon	3	0,83
zon	5	1,13
zon	6	1,12
zon	7	0,53
zon	9	1,00
bussald	0	1,77
bussald	1	1,86
bussald	2	1,78
bussald	3	1,48
bussald	4	1,00
Scale		3,35

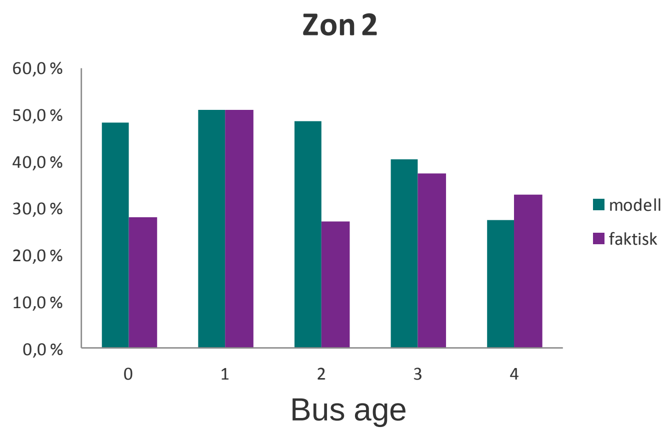
- In zon 4 (marked as 9 in the tables) the fit is ok
- There is much more data in this zon than in the others
- We may try to re-group zon, into 2,3,7 and other

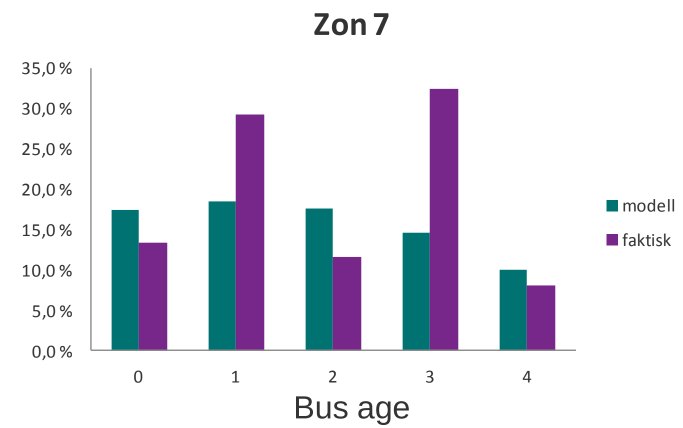
	model							
	zon	1	2	3	5	6	7	9
bus age	0	35,7 %	48,3 %	27,4 %	37,2 %	37,2 %	17,5 %	33,0 %
	1	37,7 %	51,0 %	28,9 %	39,3 %	39,2 %	18,4 %	34,9 %
	2	36,0 %	48,7 %	27,6 %	37,5 %	37,4 %	17,6 %	33,3 %
	3	29,9 %	40,4 %	22,9 %	31,1 %	31,1 %	14,6 %	27,6 %
	4	20,2 %	27,4 %	15,5 %	21,1 %	21,1 %	9,9 %	18,7 %

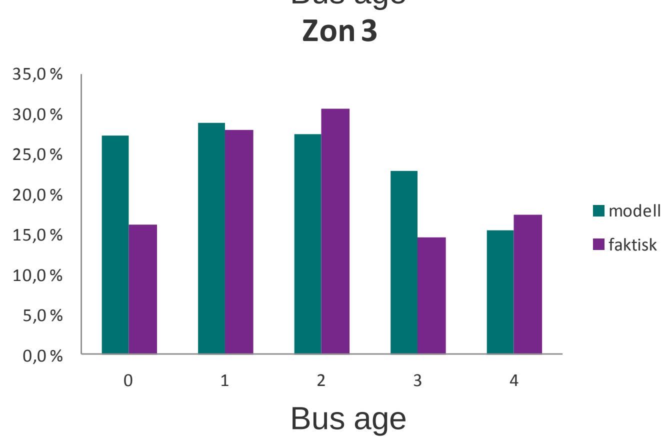
	actual							
	zon	1	2	3	5	6	7	9
bus age	0	72,6 %	28,0 %	16,2 %	32,5 %	50,9 %	13,3 %	34,3 %
	1	27,0 %	50,8 %	28,0 %	64,3 %	30,9 %	29,3 %	36,5 %
	2	32,1 %	27,1 %	30,6 %	0,0 %	24,7 %	11,7 %	38,6 %
	3	28,3 %	37,5 %	14,6 %	38,0 %	27,3 %	32,4 %	29,9 %
	4	18,5 %	32,7 %	17,3 %	21,3 %	21,9 %	8,1 %	17,7 %

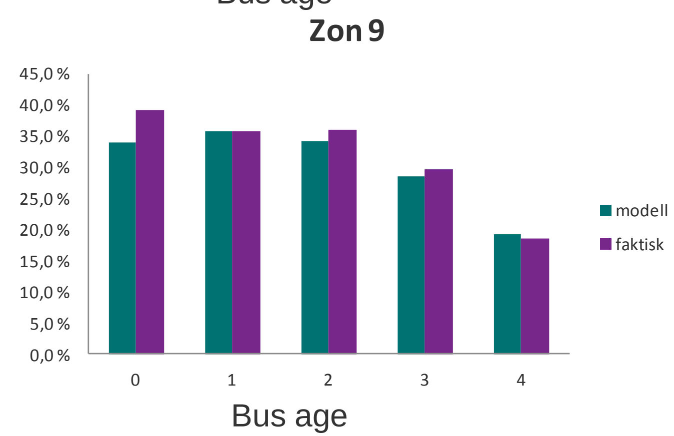
Model 2: zon regrouped

- Zon 9 (4,1,5,6) still has the best fit
- The other are better – but are they good enough?
- We try to regroup bus age as well, into 0-1, 2-3 and 4.



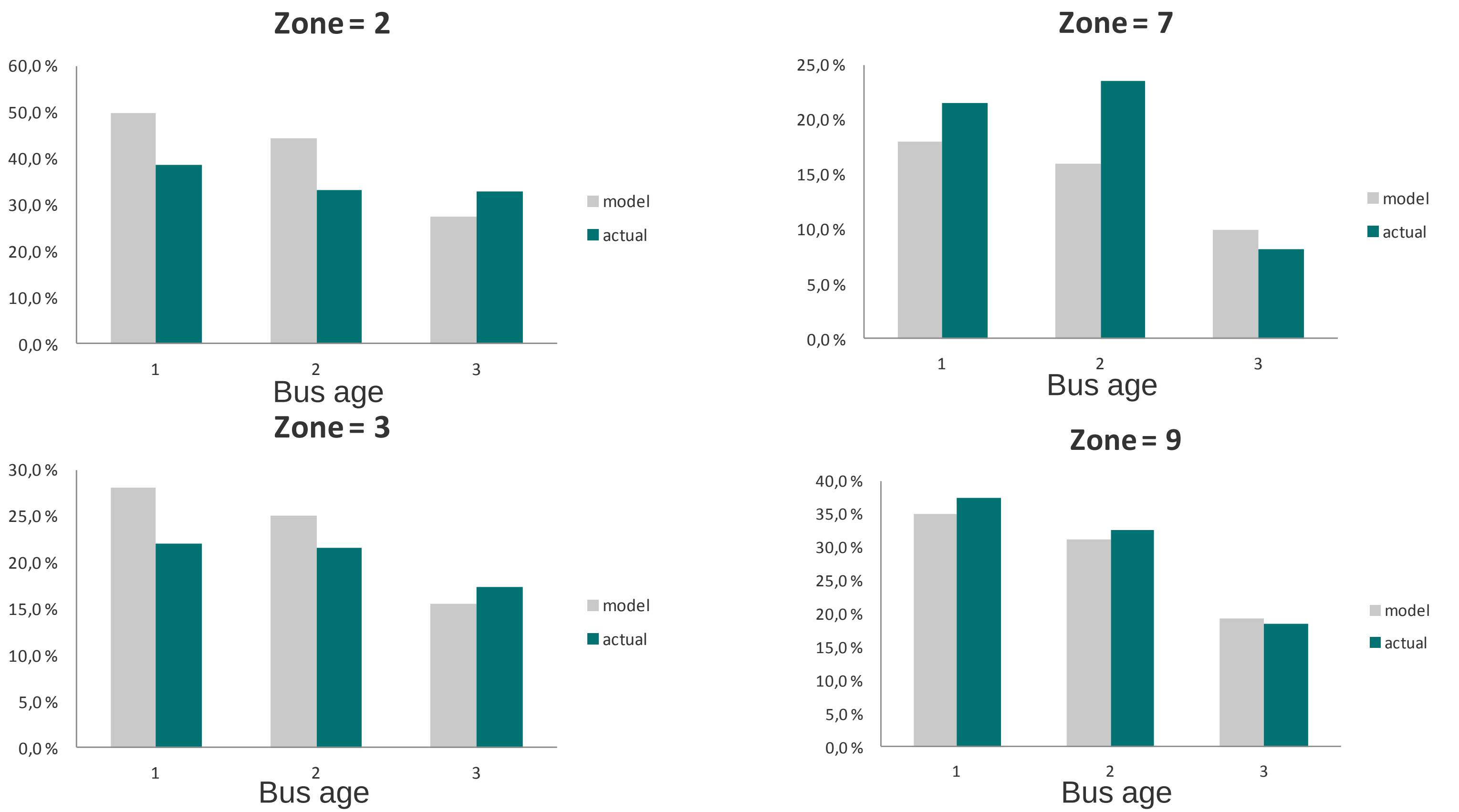






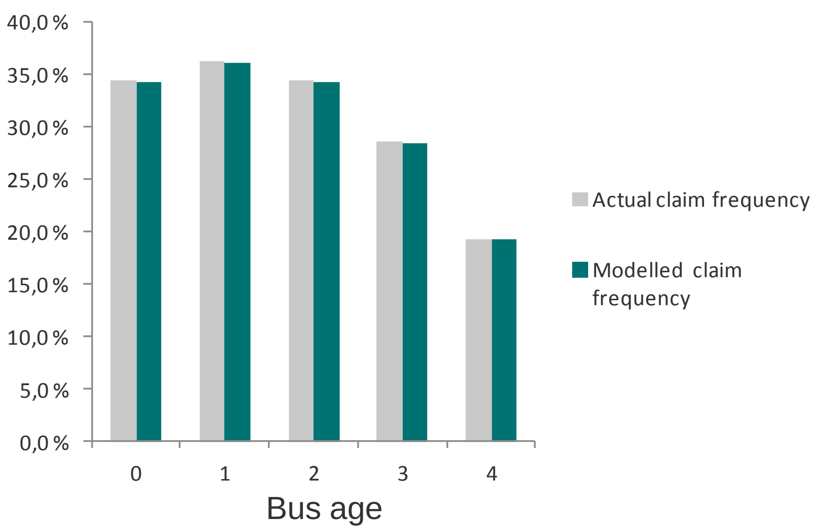
Model 3: zon and bus age regrouped

- Zon 9 (4,1,5,6) still has the best fit
- The other are still better – but are they good enough?
- May be there is not enough information in this model
- May be additional information is needed
- The final attempt for now is to skip zon and rely solely on bus age



Model 4: skip zon from the model (only bus age)

- From the graph it is seen that the fit is acceptable
- Hypothesis 1: There does not seem to be enough information in the data set to provide reliable estimates for zon
- Hypothesis 2: there is another source of information, possibly interacting with zon, that needs to be taken into account if zon is to be included in the model



Intercept		0,19
bussald	0	1,78
bussald	1	1,88
bussald	2	1,78
bussald	3	1,48
bussald	4	1,00

Limitation of the multiplicative model

- The variables in the multiplicative model are assumed to work independent of one another
- This may not be the case
- Example:
  - Auto model, Poisson regression with age and gender as explanatory variables
  - Young males drive differently (worse) than young females
  - There is a dependency between age and gender
- This is an example of an interaction between two variables
- Technically the issue can be solved by forming a new rating factor called age/gender with values
  - Young males, young females, older males, older females etc

Why is a regression model needed?

- There is not enough data to price policies individually
- What is actually happening in a regression model?
  - Regression coefficients measure the effect *ceteris paribus*, i.e. when all other variables are held constant
  - Hence, the effect of a variable can be quantified controlling for the other variables
- Why take the trouble of using a regression model?
- Why not price the policies one factor at a time?

Claim frequencies, lorry data from Länsförsäkringer (Swedish mutual)

	vehicle age		
annual milage	new	old	total
low	3,3 %	2,5 %	2,6 %
high	6,7 %	4,9 %	6,1 %
total	5,1 %	2,8 %	

- ”One factor at a time” gives 6.1%/2.6% = 2.3 as the mileage relativity
- But for each Vehicle age, the effect is close to 2.0
- ”One factor at a time” obviously overestimates the relativity – why?

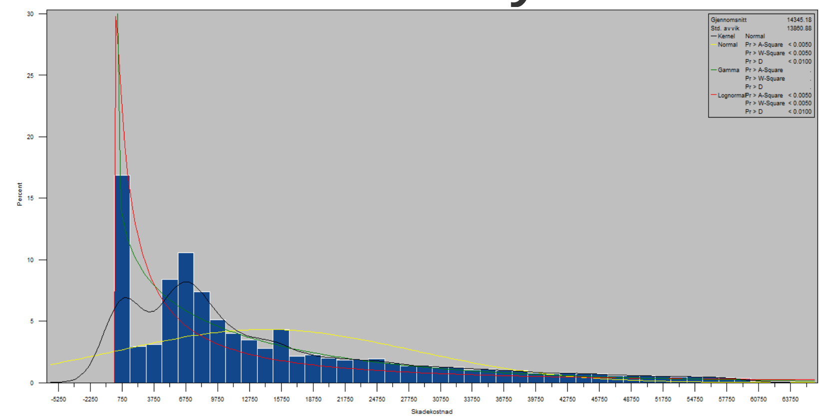
Claim frequencies, lorry data from Länsförsäkringer (Swedish mutual)

	vehicle age	
annual milage	new	old
low	47 039	190 513
high	56 455	28 612

- New vehicles have 45% of their duration in low mileage, while old vehicles have 87%
- So, the old vehicles have lower claim frequencies partly due to less exposure to risk
- This is quantified in the regression model through the mileage factor
- Conclusion: 2.3 is right for High/Low mileage if it is the only factor
- If you have both factors, 2.0 is the right relativity

Example: car insurance

- Hull coverage (i.e., damages on own vehicle in a collision or other sudden and unforeseen damage)
- Time period for parameter estimation: 2 years
- Covariates:
  - Driving length
  - Car age
  - Region of car owner
  - Tariff class
  - Bonus of insured vehicle
- Log Poisson is fitted for claim frequency
- 120 000 vehicles in the analysis



Evaluation of model

- The model is evaluated with respect to fit, result, validation of model, type 3 analysis and QQ plot
- Fit: ordinary fit measures are evaluated
- Results: parameter estimates of the models are presented
- Validation of model: the data material is split in two, independent groups. The model is calibrated (i.e., estimated) on one half and validated on the other half
- Type 3 analysis of effects: Does the fit of the model improve significantly by including the specific variable?
- QQplot:

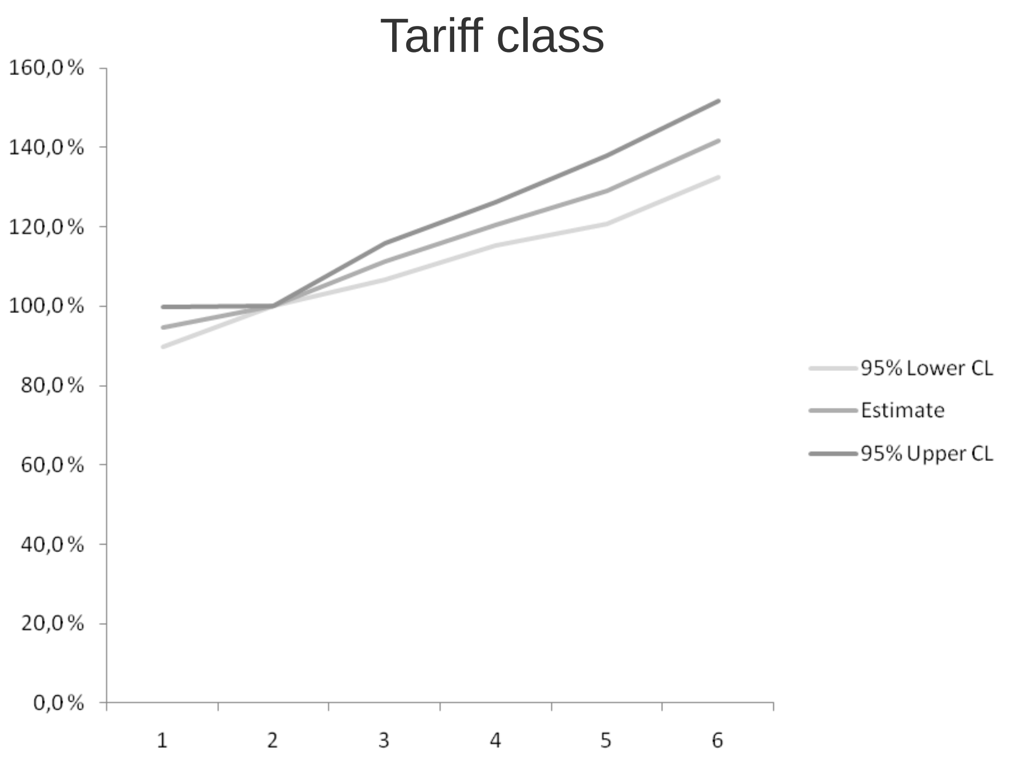
Fit interpretation

Criterion	Deg. fr.	Verdi	Value/DF
Deviance	2 365	2 337,1581	0,9882
Scaled Deviance	2 365	2 098,5720	0,8873
Pearson Chi-Square	2 365	2 633,8763	1,1137
Scaled Pearson X2	2 365	2 365,0000	1,0000
Log Likelihood	_	27 694,4040	_
Full Log Likelihood	_	- 5 078,4114	_
AIC (smaller is better)	_	10 204,8227	_
AICC (smaller is better)	_	10 205,3304	_
BIC (smaller is better)	_	10 343,5099	_

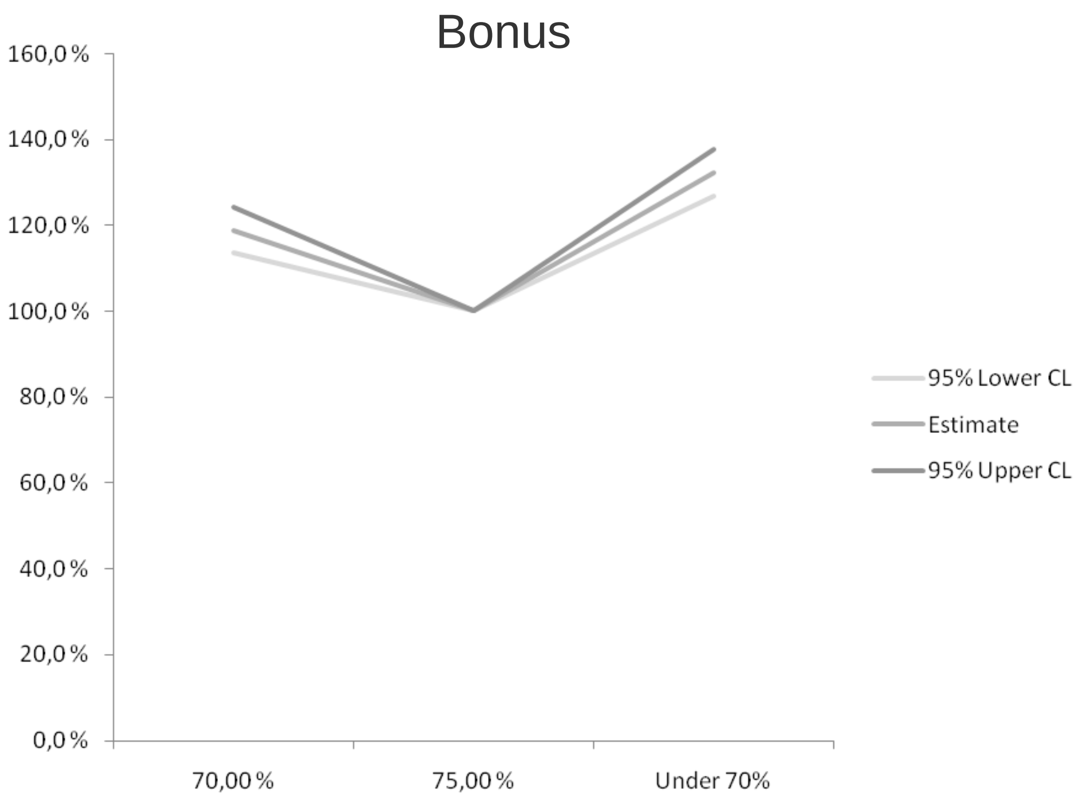
Result presentation

Variabler	Klasse	Estimat	Std. feil	Confidence	Confidence	Chi-square	Pr>Chi-sq
Intercept		- 2,0198	0,0266	- 2,0720	- 1,9676	5 757,82	<.0001
Tariff class	1	- 0,0553	0,0272	- 0,1086	- 0,0019	4,12	0,0423
Tariff class	2	0,0000	0,0000	0,0000	0,0000	.	.
Tariff class	3	0,1060	0,0209	0,0651	0,1469	25,80	<.0001
Tariff class	4	0,1873	0,0234	0,1415	0,2331	64,14	<.0001
Tariff class	5	0,2547	0,0342	0,1877	0,3216	55,58	<.0001
Tariff class	6	0,3491	0,0349	0,2807	0,4174	100,23	<.0001
Bonus	70,00 %	0,1724	0,0223	0,1287	0,2162	59,77	<.0001
Bonus	75,00 %	0,0000	0,0000	0,0000	0,0000	.	.
Bonus	Under 70%	0,2789	0,0210	0,2377	0,3201	176,04	<.0001
Region	Agder	0,0488	0,0432	- 0,0359	0,1334	1,27	0,2589
Region	Akershus Østfold	0,0000	0,0000	0,0000	0,0000	.	.
Region	Buskerud Hedmark Opplan	0,0213	0,0254	- 0,0284	0,0711	0,71	0,4007
Region	Hordaland	- 0,0393	0,0327	- 0,1033	0,0247	1,45	0,2293
Region	M R Rogaland S F	- 0,0131	0,0302	- 0,0723	0,0461	0,19	0,6644
Region	Nord	0,0487	0,0251	- 0,0006	0,0979	3,74	0,053
Region	Oslo	0,1424	0,0259	0,0917	0,1931	30,33	<.0001
Region	Telemark Vestfold	0,0230	0,0312	- 0,0380	0,0841	0,55	0,4596
Driving Length	8000	- 0,1076	0,0252	- 0,1570	- 0,0583	18,27	<.0001
Driving Length	12000	0,0000	0,0000	0,0000	0,0000	.	.
Driving Length	16000	0,1181	0,0214	0,0761	0,1601	30,41	<.0001
Driving Length	20000	0,2487	0,0237	0,2022	0,2951	110,08	<.0001
Driving Length	25000	0,4166	0,0336	0,3508	0,4824	153,86	<.0001
Driving Length	30000	0,5687	0,0398	0,4906	0,6467	204,04	<.0001
Driving Length	99999	0,8168	0,0500	0,7188	0,9149	266,54	<.0001
Car age	1	- 0,0136	0,0240	- 0,0607	0,0335	0,32	0,5715
Car age	2	0,0000	0,0000	0,0000	0,0000	.	.
Car age	3	- 0,0638	0,0177	- 0,0986	- 0,0290	12,94	0,0003
Car age	4	- 0,0800	0,0386	- 0,1400	- 0,0400	9,38	0,0022

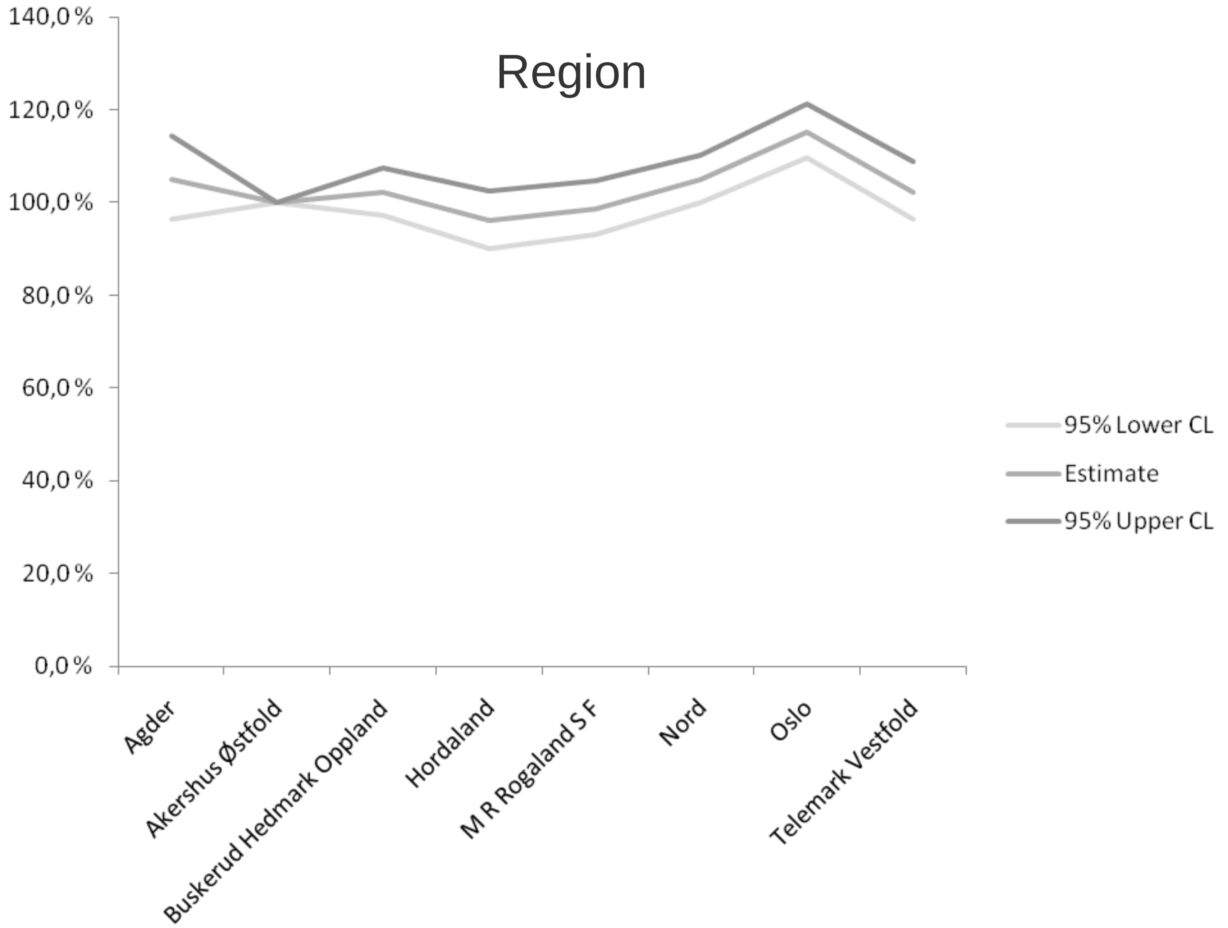
Result presentation



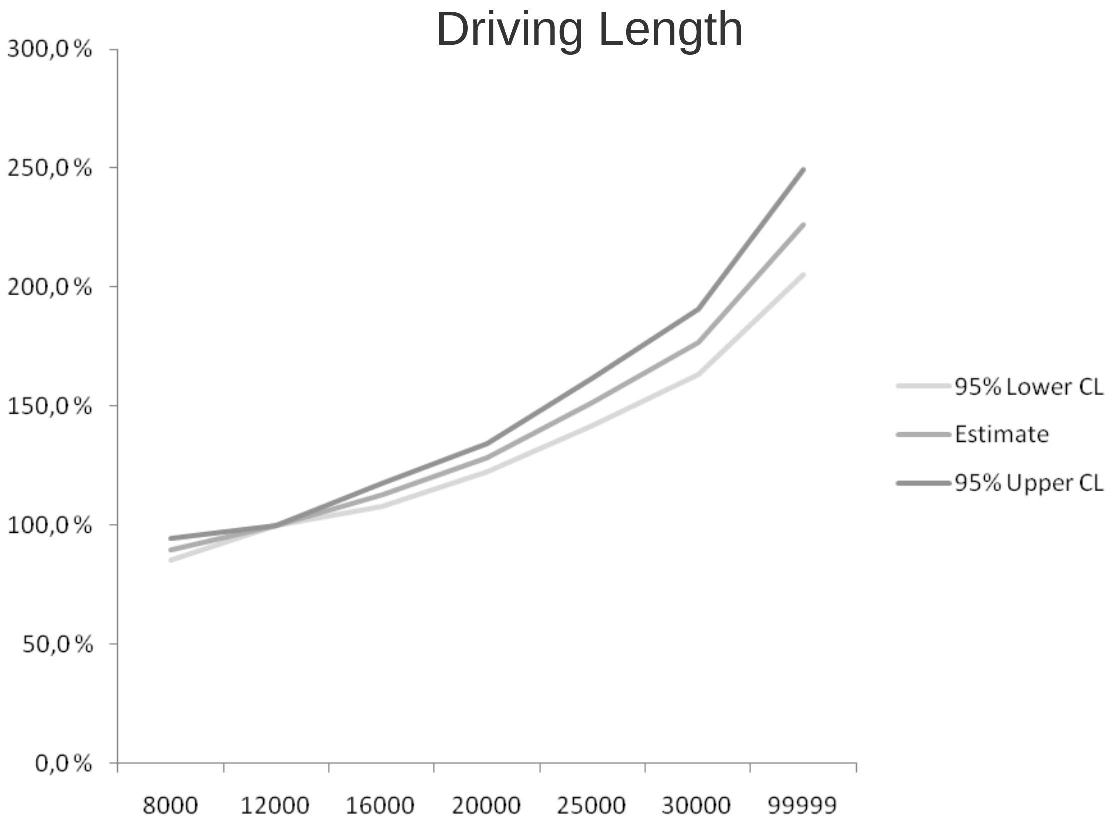
Result presentation



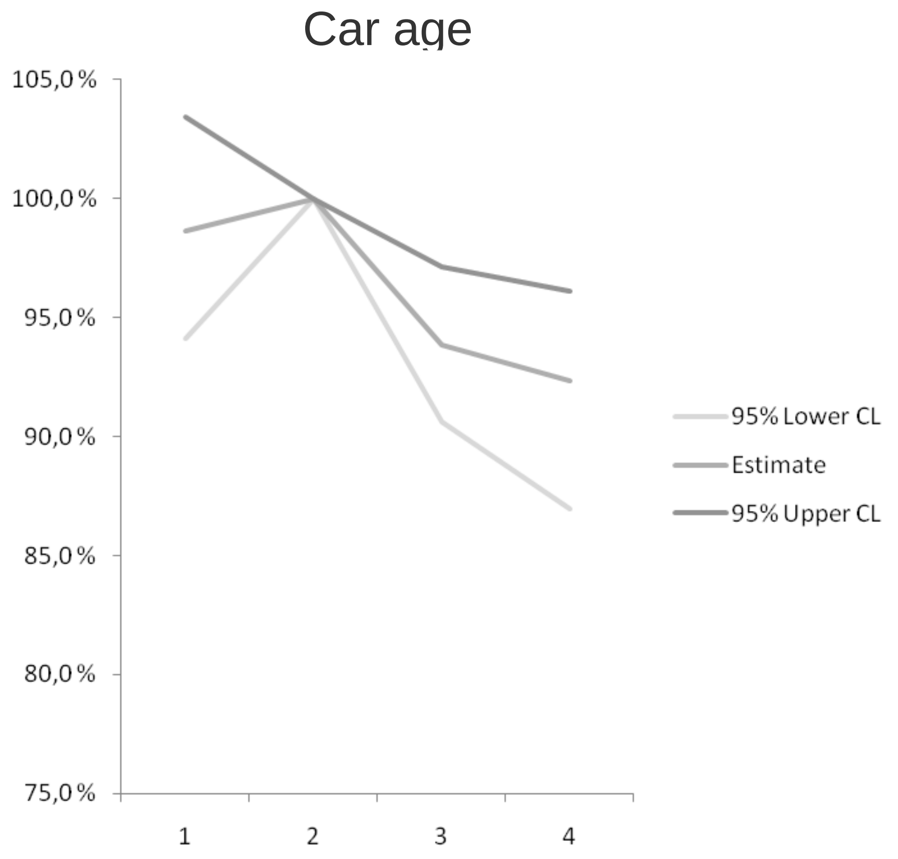
Result presentation



Result presentation



Result presentation



Validation

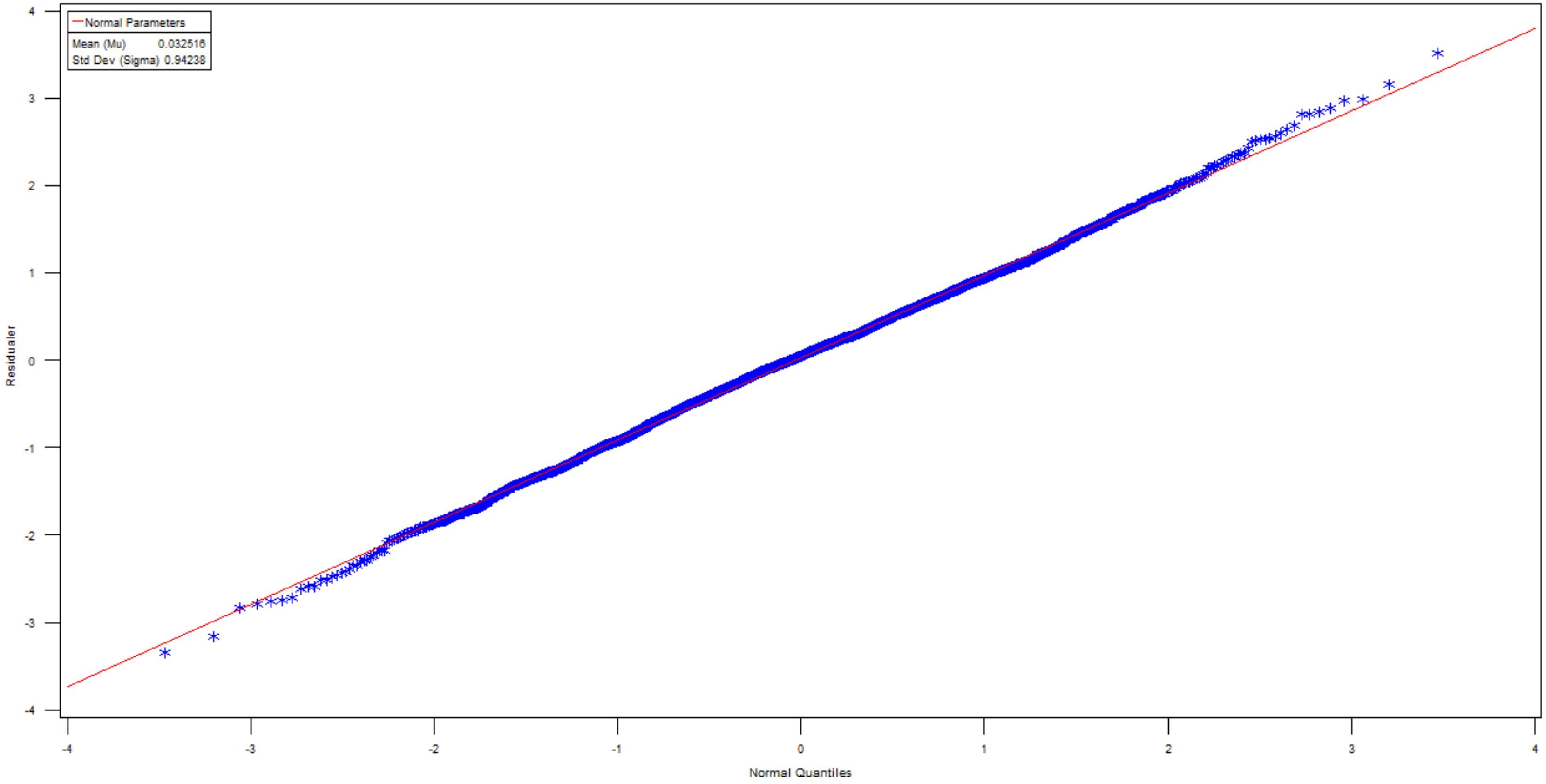
Variables	Class	Model	Portfolio	Diff.
Tariff class	Total	19 332	18 284	5,73
Tariff class	1	2 360	2 138	10,37
Tariff class	2	5 059	4 921	2,81
Tariff class	3	5 586	5 426	2,95
Tariff class	4	3 686	3 442	7,08
Tariff class	5	1 367	1 227	11,38
Tariff class	6	1 274	1 130	12,77
Bonus	Total	19 332	18 284	5,73
Bonus	70,00 %	3 103	2 851	8,83
Bonus	75,00 %	12 696	12 116	4,79
Bonus	Under 70%	3 533	3 317	6,53
Region	Total	19 332	18 284	5,73
Region	Agder	805	713	12,87
Region	Akershus Østfold	4 356	4 335	0,47
Region	Buskerud Hedmark Oppland	3 078	2 866	7,41
Region	Hordaland	1 497	1 432	4,53
Region	M R Rogaland S F	1 839	1 672	9,99
Region	Nord	3 163	2 920	8,33
Region	Oslo	2 917	2 773	5,21
Region	Telemark Vestfold	1 677	1 573	6,62
Driving Length	Total	19 332	18 284	5,73
Driving Length	8000	2 793	2 752	1,48
Driving Length	12000	5 496	5 350	2,73
Driving Length	16000	4 642	4 480	3,62
Driving Length	20000	3 427	3 247	5,54
Driving Length	25000	1 376	1 193	15,37
Driving Length	30000	995	813	22,37
Driving Length	99999	603	449	34,33
Car age	Total	19 332	18 284	5,73
Car age	<= 5 år	2 710	2 386	13,59
Car age	5-10år	9 255	9 052	2,24
Car age	10-15år	6 299	6 032	4,43
Car age	>15 år	1 068	814	31,17

Type 3 analysis

Type 3 analysis of effects: Does the fit of the model improve significantly by including the specific variable?

Source	Num DF	Den DF	F Value	Pr > F	Chi-square	Pr>Chi-sq	Method
Tariff class	5	2 365	38,43	<.0001	192,17	<.0001	LR
Bonus	2	2 365	97,09	<.0001	194,18	<.0001	LR
Region	7	2 365	6,35	<.0001	44,48	<.0001	LR
Driving Length	6	2 365	97,51	<.0001	585,08	<.0001	LR
Car age	3	2 365	9,23	<.0001	27,69	<.0001	LR

Type 3 analysis



Some repetition of generalized linear models (GLMs)

Exponential dispersion Models (EDMs)

- Frequency function f_{Y_i} (either density or probability function)

$$f_{Y_i}(y_i;\theta_i,\varphi)=\exp\{\frac{y_i\theta_i-b(\theta_i)}{\varphi/w_i}+c(y_i,\varphi,w_i)\}$$

For yi in the support, else $f_{Y_i}=0$.

- c() is a function not depending on θ_i
- $\varphi>0, w_i>0$ or $w_i\geq 0$
- $b(\theta_i)$ twice differentiable function
- b' has an inverse
- The set of possible θ_i is assumed to be open

Claim frequency

- Claim frequency $Y_i=X_i/T_i$ where T_i is duration
- Number of claims assumed Poisson with $E(X_i)=T_i\mu_i$

$$f_{Y_i}(y_i;\theta_i,\varphi)=P(Y_i=y_i)=P(X_i=T_iy_i)=e^{-T_i\mu_i}\frac{(T_i\mu_i)^{T_iy_i}}{(T_iy_i)!}$$

$$=\exp\{T_i[y_i\log(\mu_i)-\mu_i]\}+c\}$$

- Let $\theta_i=\log(\mu_i)$
- Then

$$f_{Y_i}(y_i;\theta_i)=\exp\{T_i(y_i\theta_i-e^{\theta_i})+c\}$$

- EDM with $\varphi=1,\ b(\theta_i)=e^{\theta_i}\quad -\infty<\theta_i<\infty$

Note that an EDM...

- ...is not a parametric family of distributions (like Normal, Poisson)
- ...is rather a class of different such families
- The function b() speficies which family we have
- The idea is to derive general results for all families within the class – and use for all

Expectation and variance

- By using cumulant/moment-generating functions, it can be shown (see McCullagh and Nelder (1989)) that for an EDM

$$\mu_i \equiv E(Y_i) = b'(\theta_i)$$
$$\mathrm{var}(Y_i) = b''(\theta_i)\varphi / T_i$$

- This is why b() is called the *cumulant function*

The variance function

- Recall that $\theta_i = b'^{-1}(\mu_i)$ is assumed to exist
- Hence $b''(\theta_i) = b''(b'^{-1}(\mu_i))$
- The variance function is defined by $v(\mu_i) = b''(b'^{-1}(\mu_i))$
- Hence $\mathrm{var}(Y_i) = \varphi \dfrac{v(\mu_i)}{w_i}$

Common variance functions

Distribution	Normal	Poisson	Gamma	Binomial
$v(\mu_i)$	1	μ	μ^2	$\mu(1-\mu)$

Note: Gamma EDM has std deviation proportional to μ , which is much more realistic than constant (Normal)

Theorem

Within the EDM class, a family of probability distributions is uniquely characterized by its variance function

Proof by professor Bent Jørgensen, Odense

Scale invariance

- Let $c>0$
- If cY belongs to same distribution family as Y, then distribution is *scale invariant*
- Example: claim cost should follow the same distribution in NOK, SEK or EURO

Tweedie Models

- If an EDM is scale invariant then it has variance function

$$v(\mu_i) = \mu^{\mathrm{p}}$$

- This is also proved by Jørgensen
- This defines the Tweedie subclass of GLMs
- In pricing, such models can be useful

Overview of Tweedie Models

	Type	Name	Key ratio
p<0	Continuous	-	-
p=0	Continuous	Normal	-
0<p<1	Non-existing	-	-
p=1	Discrete	Poisson	Claim frequency
1<p<2	Mixed, non-negative	Compound Po	Pure premium
p=2	Continuous, positive	Gamma	Claim severity
2<p<3	Continuous, positive	-	(claim severity)
p=3	Continuous, positive	Inverse Norma	(claim severity)
p>3	Continuous, positive	-	(claim severity)

Link functions

- A general link function g()

$$g(\mu_i) = \sum_{j=1}^{r} x_{ij}\beta_j, \;\; i = 1,\ldots,n$$

- Linear regression: $g(\mu_i) = \mu_i$ identity link
- Multiplicative model: $g(\mu_i) = \log \mu_i$ log link
- Logistic regression: $g(\mu_i) = \log(\frac{\mu_i}{1-\mu_i})$ logit link

Summary

Generalized linear models:

- Yi follows an EDM: $\mathrm{var}(Y_i) = \varphi v(\mu_i) / w_i$
- Mean satisfies $g(\mu_i) = \sum_j x_{ij} \beta_j$

Multiplicative Tweedie models:

- Yi Tweedie EDM: $\mathrm{var}(Y_i) = \varphi \mu_i^{\,p} / w_i, \; p \geq 1$
- Mean satisfies $\log(\mu_i) = \sum_j x_{ij} \beta_j$