Europa-Universität Viadrina Frankfurt (Oder)

Fast initial response für EWMA-Kontrollkarten

Sven Knoth

Lehrstuhl für Quantitative Methoden, insbesondere Statistik

13. Juni 2003

Traditionelle EWMA-Kontrollkarte für die Lage

(mit konstanten Stoppgrenzen)

$$X_i \sim \mathcal{N}(\mu_i, \sigma^2 = 1)\,, \quad \mu_i = \begin{cases} 0 & , \; i < m \\ \delta\;(=1) & , \; i \geq m \end{cases}\,, \quad i = 1, 2, \ldots$$

$$Z_0 = z_0 = 0\,, \quad Z_n \underset{n \geq 1}{=} (1-\lambda)\, Z_{n-1} + \lambda\, X_n\,, \quad \lambda \in (0, 1]\,,$$

$$E_m(Z_n) \underset{n<m}{=} 0 \quad \Big(E_\infty(Z_n) = 0 \text{ für alle } n, \text{ for } z_0 \neq 0 \text{ wenigstens } \lim_{n\to\infty} E_\infty(Z_n) = 0\Big)\,,$$

$$Var_m(Z_n) = \frac{2}{2-\lambda}\left(1 - [1-\lambda]^{2\,n}\right) \to \frac{2}{2-\lambda} \text{ for } n \to \infty\,,$$

$$L = \inf\left\{ n \in \mathbb{N} : |Z_n| > c\sqrt{\frac{2}{2-\lambda}} \right\}.$$

Fast Initial Response (FIR)

Ziel: Verbessere Verhalten der Kontrollkarte zum Verfahrensstart!

z. B.

- nach einer Prozeßadjustierung in der Qualitätskontrolle,
- nach Modellneuschätzung beim Monitoring von Finanzdaten,
- bei vor Verfahrensstart vermutetem *change point*,
- Verringerung des Einflusses des Startwertes (z_0),
- ...

Fast Initial Response – Literaturstellen

- Erstmalig betrachtet von Lucas/Crosier (1982) für CUSUM.

- Lucas/Saccucci (1990 ... 1988) `Markov-Kette`

  5.1 FIR Feature ... A disadvantage of this approach is that it requires two separate EWMA's ... A detailed discussion of this approach will be given in a future article.

  B.1 FIR Feature ... The FIR feature requires the simultaneous implementation of two one-sided EWMA ... We plan to implement the exact method for calculating FIR ARL's in our future work.

- Rhoads/Montgomery/Mastrangelo (1996), `Monte-Carlo`

  Fast initial response scheme for exponentially weighted moving average control chart.

- Steiner (1999), `Markov-Kette`

  EWMA control charts with time-varying control limits and fast initial response.

- Montgomery (2001): $\frac{1}{2}$ Seite in *Introduction to Statistical Quality Control*, 4^{th} *ed.*

FIR als Ausweg für Trägheitsproblem

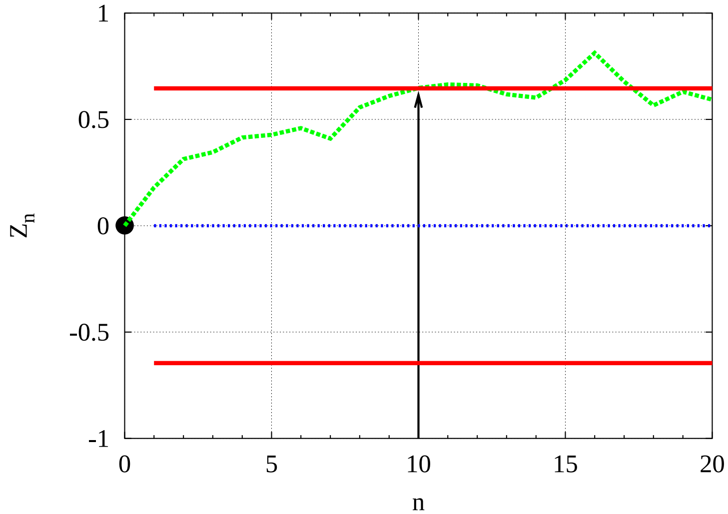
Problem: träges Verhalten der traditionellen EWMA-Karte für zeitige *change points* (kleine m), d. h. bei kleinem λ starker Einfluß des Startwertes z_0

$$Var_m(Z_n) = \frac{2}{2-\lambda}\left(1 - [1-\lambda]^{2n}\right) \ll \frac{2}{2-\lambda} = \lim_{n\to\infty} Var_m(Z_n)$$

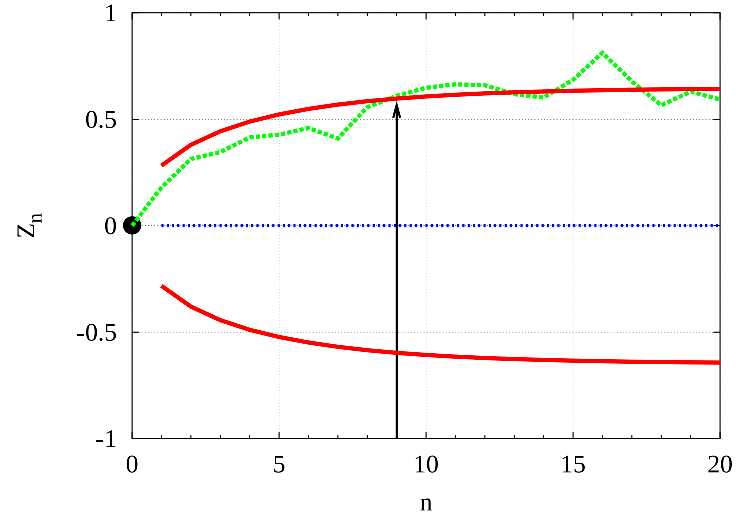
Illustrierende Beispiele: $\lambda = 0.1$, $m = 1$, $E_\infty(L) = 500$

FIR EWMA-Beispiele I – $m = 1$

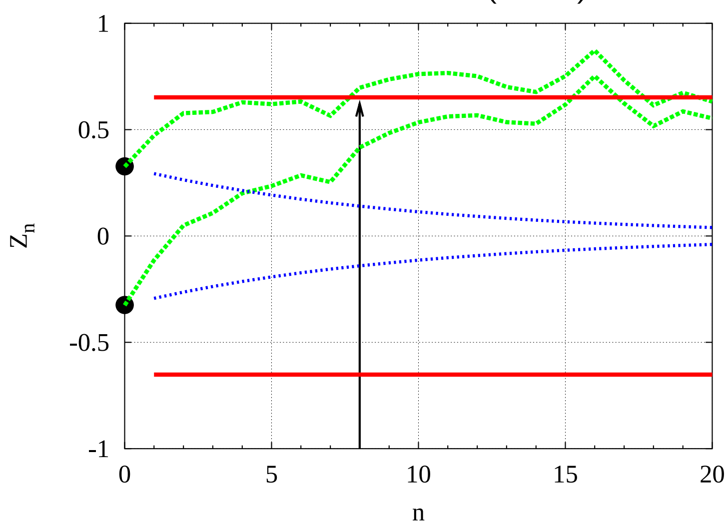
Roberts (1959), Wortham (1972), Ho (1978)



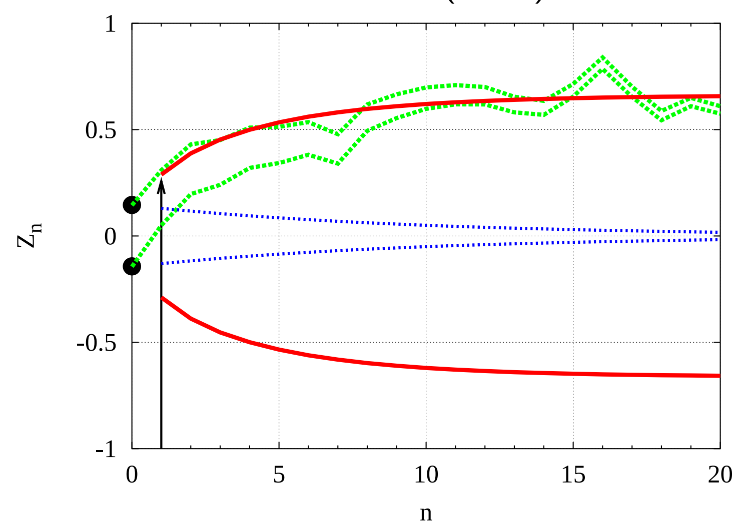
Roberts (1959), Chandrasekaran et al. (1995)



Lucas and Saccucci (1990)

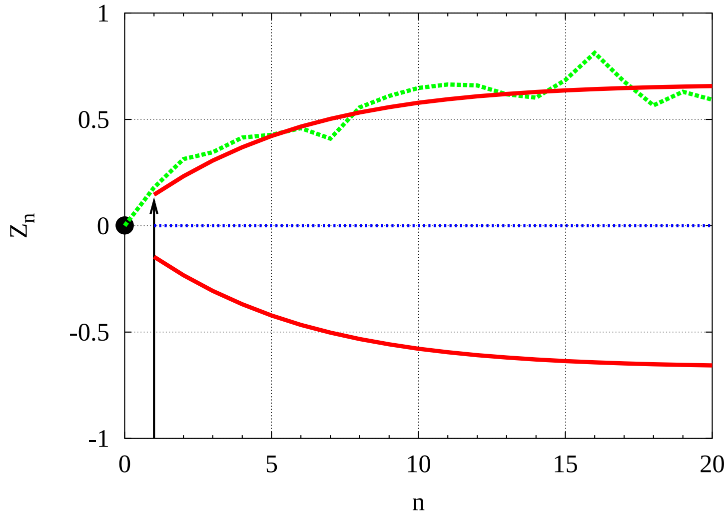


Rhoads et al. (1996)

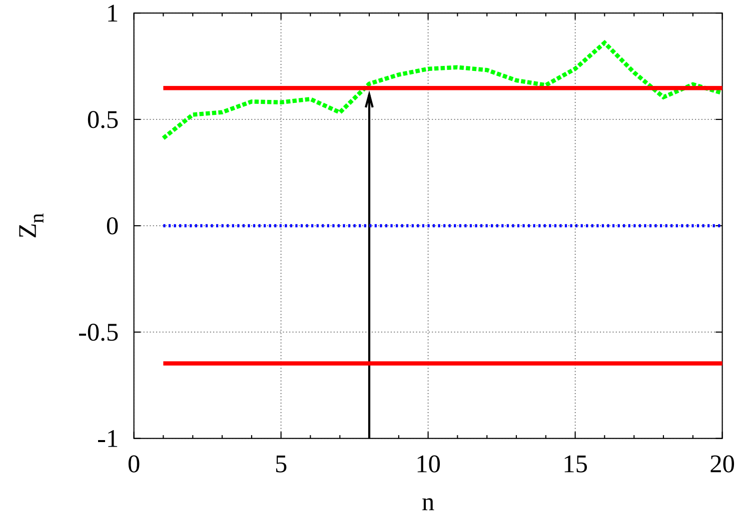


FIR EWMA-Beispiele II – $m = 1$

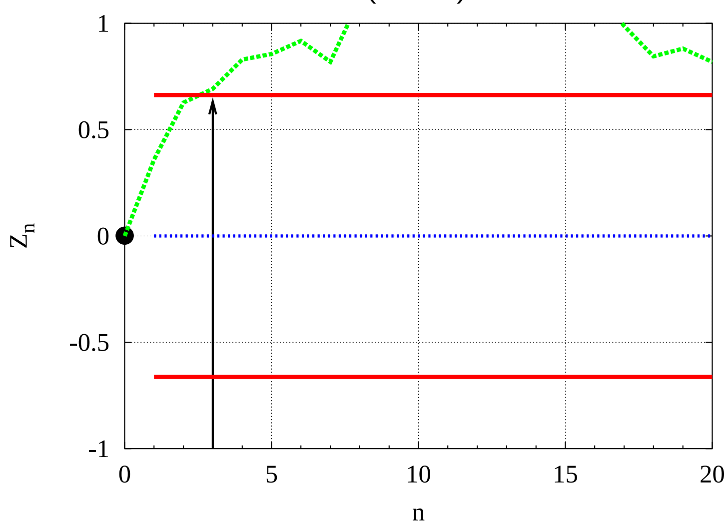
Steiner (1999)



(Knoth 2002)



Fink (1994)



?

?

Formeln

$$Z_0 = 0 \ , \ Z_n \underset{n \geq 1}{=} (1-\lambda)\, Z_{n-1} + \lambda\, X_n \ , \ c^* = c \sqrt{\frac{2}{2-\lambda}} \ , \ l_n = \sqrt{1-(1-\lambda)^{2n}}$$

Alarm, wenn

(fcl) $|Z_n| > c^* \,,$

(vacl) $|Z_n| > c^*\, l_n \,,$

(fir) $\max\left\{|Z_n^u|, |Z_n^l|\right\} > c^* \ , \ Z_0^u = -Z_0^l = \frac{c^*}{2} \,,$ `Lucas/Saccucci 1990`

(fvacl) $\max\left\{|Z_n^u|, |Z_n^l|\right\} > c^*\, l_n \ , \ Z_0^u = -Z_0^l = l_1 \cdot \frac{c^*}{2} \,,$ `Rhoads et al. 1996`

(fadj) $|Z_n| > c^*\, l_n \left(1-(1-f)^{1+a(n-1)}\right) \ , \ f = 0.5, \ a = 0.3 \,,$ `Steiner 1999`

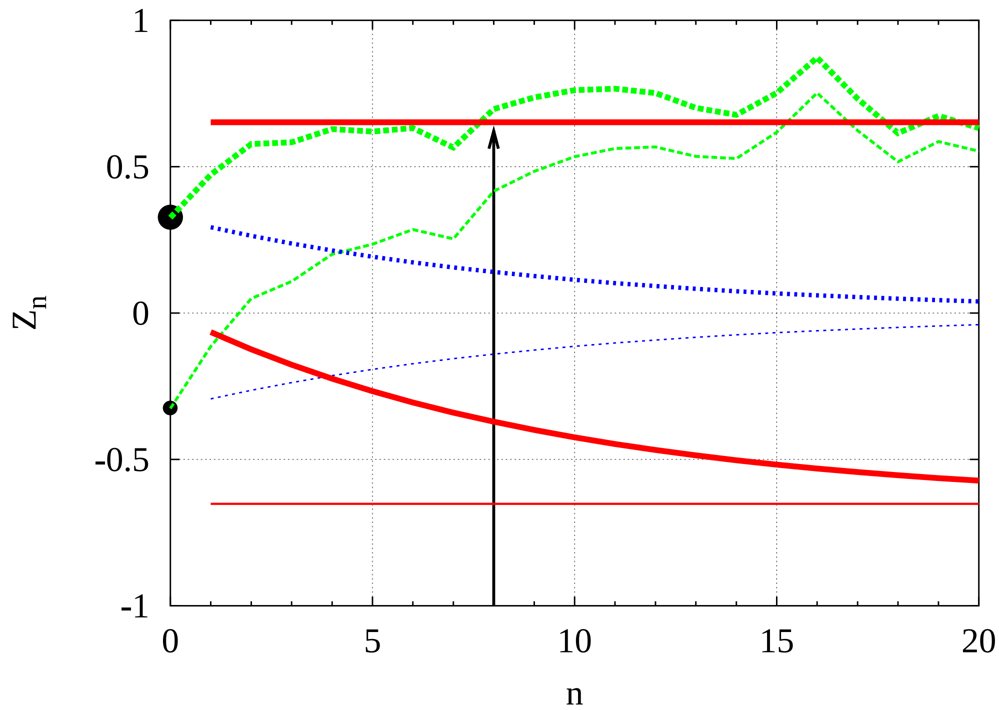
(stat) $|Z_n| > c^* \ , \ Z_1 = X_1 \sqrt{2/(2-\lambda)}\,, \ Var_m(Z_n) = \frac{2}{2-\lambda} \,,$ `(Knoth 2002)`

(switch) $|Z_n| > c^* \ , \ \lambda = \begin{cases} \lambda_0 & , \ n \leq n_1 \\ \lambda_1 < \lambda_0 & , \ n > n_1 \end{cases} \ , \ n_1 = 10, \lambda_1 = \lambda_0/2 \,.$ `Fink 1994`

FIR nach Lucas/Saccucci benötigt zwei Karten?!

$$Z_0^u = -Z_0^l = \frac{c^*}{2}\,,$$

$$Z_n^l = Z_n^u - \left(Z_0^u - Z_0^l\right)(1-\lambda)^n$$
$$= Z_n^u - c^*\,(1-\lambda)^n$$



Deshalb ist, $\left(L = \inf\left\{n \in \mathbb{N} : \ldots\right\}\right)$

$$\max\left\{|Z_n^u|, |Z_n^l|\right\} > c^* \Leftrightarrow \max\left\{Z_n^u, -Z_n^l\right\} > c^*$$
$$\Leftrightarrow Z_n^u \notin \left[-\left(1-(1-\lambda)^n\right)c^*, c^*\right]$$

Bem.: Analoges gilt für FIR nach Rhoads et al.

Berechnung von $E_m(L)$, des kritischen Wertes c und ...

- Für konstante Grenzen (`fcl, stat`) gibt es "Dutzende" Methoden, für `switch` ist eine einfache Anpassung möglich und für `fir` gibt es zumindest den 2-dim. Zugang (Lucas/Saccucci).

- Für EWMA-Kontrollkarten mit variierenden Grenzen findet man:
  - Markov-Ketten, z. B. in Chandrasekaran et al. (1995) und Steiner (1999),
  - Monte-Carlo, z. B. in Rhoads et al. (1996),
  - Dichterekursionen in Margavio, Conerly, Woodall, Drake (1995)

  Blinder-Alarm-Rate $r_n = P_\infty(L = n | L \geq n)$ für `fcl, vacl` und eine spezielle Karte mit konstanter Rate

Dichterekursionen I

Eine mehr allgemeine Schreibweise ist $L = \inf \left\{ n \in \mathbb{N} : Z_n \notin [a_n, b_n] \right\}$.

Betrachten die folgende Rekursion:

$$f_1(z; z_0) = \begin{cases} \frac{1}{\lambda}\,\phi_\mu\!\left(\frac{z-(1-\lambda)z_0}{\lambda}\right) & ,\ a_1 \le z \le b_1 \\ 0 & ,\ \text{sonst} \end{cases} ,$$

$$f_n(z; z_0) = \begin{cases} \displaystyle\int_{a_{n-1}}^{b_{n-1}} f_{n-1}(\tilde{z}; z_0)\,\frac{1}{\lambda}\,\phi_\mu\!\left(\frac{z-(1-\lambda)\tilde{z}}{\lambda}\right) d\tilde{z} & ,\ a_n \le z \le b_n \\ 0 & ,\ \text{sonst} \end{cases} .$$

Dann ist $f_n^*(z; z_0) = f_n(z; z_0) / \int_{a_n}^{b_n} f_n(\tilde{z}; z_0)\, d\tilde{z}$

eine Dichte der EWMA-Statistik Z_n bedingt auf $L > n$ und $Z_0 = z_0$,

wobei der Nenner gerade gleich $P(L > n)$ ist.

Dichterekursionen II

Mittels $P(L > n)$ ergeben sich

`zero-state ARL` $$E_{1(\infty)}(L) = \sum_{n=0}^{\infty} P_{1(\infty)}(L > n)\,,$$

$$D_m = E_m(L - m + 1 | L \geq m) = \frac{\sum_{n=m}^{\infty} P_m(L > n)}{P_m(L > m-1)}\,,$$

`steady-state ARL` $$D = \lim_{m\to\infty} D_m \approx D_{m_0}\;,\; m_0 = 100\,,$$

$$P_m(L = n)\;,\; P_m(L = n + m | L \geq m)\;,\; L_\alpha\;,\; r_n\;,\; \ldots\;,$$

c als Lösung von $E_\infty(L) = A = 500\,.$

Dichterekursionen III

Wenn $\lim_{n\to\infty} a_n = a$ und $\lim_{n\to\infty} b_n = b$,

dann verhalten sich die Tails der Stoppzeit L wie die einer geometrischen Verteilung,

d. h. $P_m(L > n + j) \approx \varrho^j \, P_m(L > n)$

mit ϱ als Eigenwert mit dem größten Betrag bzgl. des Grenzübergangskernes.

$\rightsquigarrow$ nutze Quadraturen bis zu einem gewissen n_0 and dann den geometrischen Tail,

z. B. $E_\infty(L) \approx \sum_{i=0}^{n_0-1} P(L > n) + \frac{P(L > n_0)}{1 - \tilde{\varrho}}, \qquad \tilde{\varrho} = \frac{P(L>n_0)}{P(L>n_0-1)}.$

Bem.:

1. Für homogene Kerne nutzten Woodall (1983), Waldmann (1986) und Gan (1991) dieselbe Idee (Ausnahmen sind D_m und D).
2. Es ist unklar, wie in Margavio et al. (1995) die variierenden Grenzen in der Quadraturfolge behandelt worden sind und warum z. B. Steiner (1999) diese Arbeit weder nutzte noch zitierte.

Dichterekursionen IV

$$-a_n = b_n = c^* \qquad \texttt{fcl, stat, switch}\,,$$

$$-a_n = b_n = c^* \sqrt{1-(1-\lambda)^{2n}} = c^*\, l_n \qquad \texttt{vacl}\,,$$

$$a_n = -\big(1-(1-\lambda)^n\big)\, c^*\,,\ b_n = c^* \qquad \texttt{fir}\,,$$

$$a_n = -\big(1-(1-\lambda)^n\big)\, c^*\, l_n\,,\ b_n = c^*\, l_n \qquad \texttt{fvacl}\,,$$

$$-a_n = b_n = c^*\, l_n \left(1-(1-f)^{1+a(n-1)}\right) \qquad \texttt{fadj}\,.$$

$$\lim_{n\to\infty} -a_n = \lim_{n\to\infty} b_n = c^*\,.$$

$$r_n = (b_n - a_n)/(2c^*)\ ,\ d_n = (a_n + b_n)/2\,,$$

$$\begin{aligned} M_n(\tilde{z}, z) &= r_n M(d_{n-1} + r_{n-1}\tilde{z}, d_n + r_n z) \\ &= \frac{r_n}{\lambda}\phi_\mu\left(\frac{d_n + r_n z - (1-\lambda)(d_{n-1} + r_{n-1}\tilde{z})}{\lambda}\right)\ . \end{aligned}$$

$\rightsquigarrow$ glatter und positiver Übergangskern auf $[-c^*, c^*] \times [-c^*, c^*]$,

d. h. nutze Quadraturknoten unabhängig von n und erhalte gleichzeitig akkurate Quadraturresultate.

Kritische Werte

unter-Kontrolle-ARL, d. h. $E_\infty(L) = 500$, $\lambda = 0.1$

EWMA-Karte	fcl	vacl	fir	fvacl	fadj	stat	switch
c	2.8143	2.8239	2.8415	2.8858	2.9131	2.8215	2.8879
$\widehat{E_\infty(L)}$, 10^9 rep.	499.99	500.02	500.01	499.92	500.02	500.01	499.94
se	.016	.016	.017	.017	.020	.016	.019
`Dichterekursion`	`499.99`	`500.04`	`499.99`	`499.93`	`500.04`	`499.99`	`499.97`

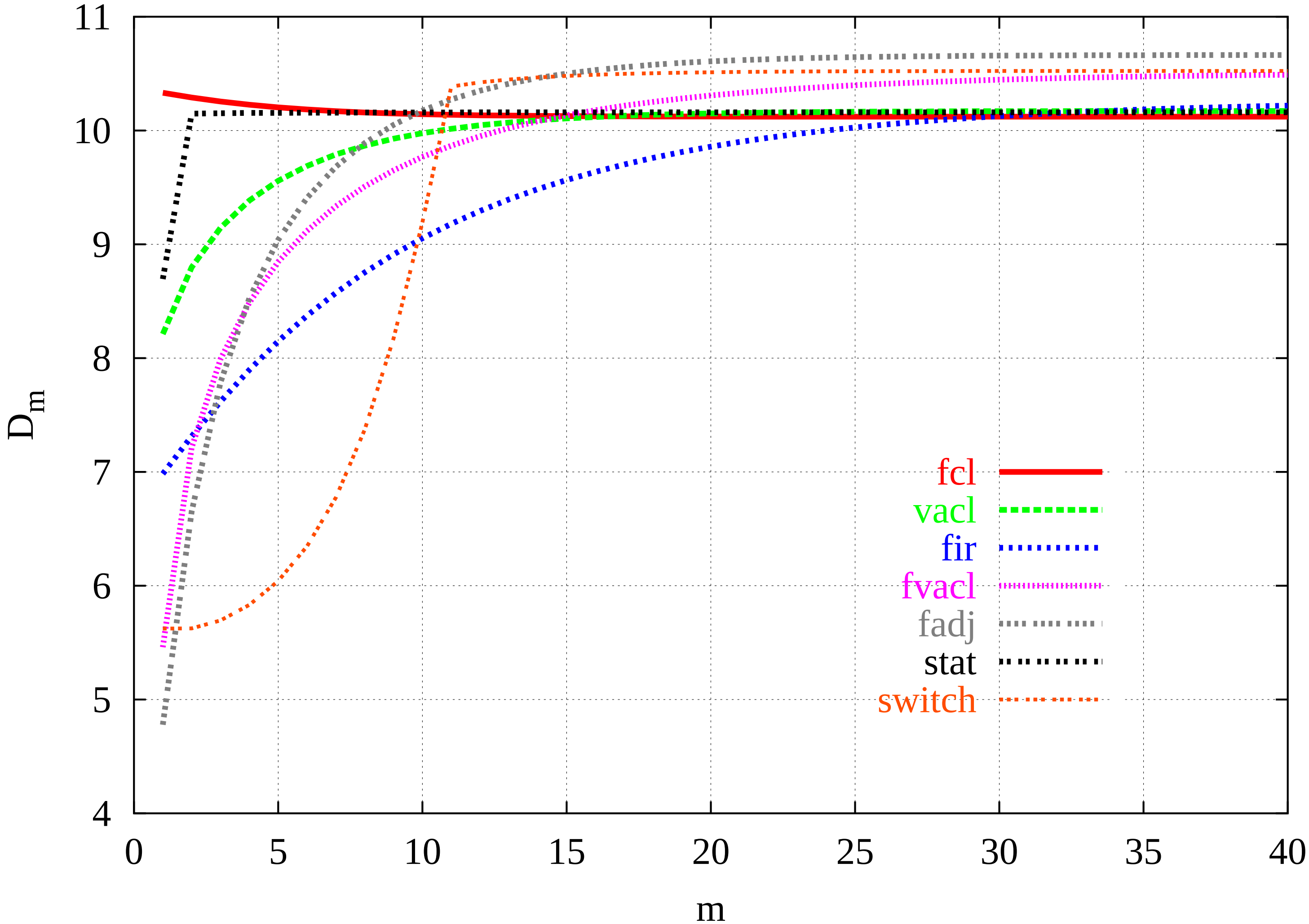
Zero-state und steady-state ARLs $E_1(L)$ und D

$\lambda = 0.1$, unter-Kontrolle-ARL $E_\infty(L) = 500$, Ergebnisse als $\begin{pmatrix} E_1(L) \\ D \end{pmatrix}$

	EWMA-Kontrollkarte						
δ	fcl	vacl	fir	fvacl	fadj	stat	switch
0.5	$\overline{\mathbf{31.3}}$	28.8	24.8	22.9	21.6	29.3	$\underline{\mathbf{20.8}}$
	$\underline{\mathbf{30.6}}$	30.9	31.4	32.8	$\overline{\mathbf{33.6}}$	30.8	32.8
1.0	$\overline{\mathbf{10.3}}$	8.21	6.98	5.46	$\underline{\mathbf{4.78}}$	8.69	5.62
	$\underline{\mathbf{10.1}}$	10.2	10.3	10.5	$\overline{\mathbf{10.7}}$	10.2	10.5
1.5	$\overline{\mathbf{6.08}}$	4.17	3.90	2.52	$\underline{\mathbf{2.19}}$	4.56	3.36
	$\underline{\mathbf{5.99}}$	6.01	6.06	6.17	$\overline{\mathbf{6.24}}$	6.01	6.18
2.0	$\overline{\mathbf{4.36}}$	2.66	2.75	1.60	$\underline{\mathbf{1.45}}$	2.91	2.47
	$\underline{\mathbf{4.31}}$	4.32	4.35	4.42	$\overline{\mathbf{4.47}}$	4.32	4.43
3.0	$\overline{\mathbf{2.87}}$	1.51	1.81	1.09	$\underline{\mathbf{1.07}}$	1.57	1.68
	$\underline{\mathbf{2.85}}$	2.86	2.87	2.91	$\overline{\mathbf{2.94}}$	2.85	2.92

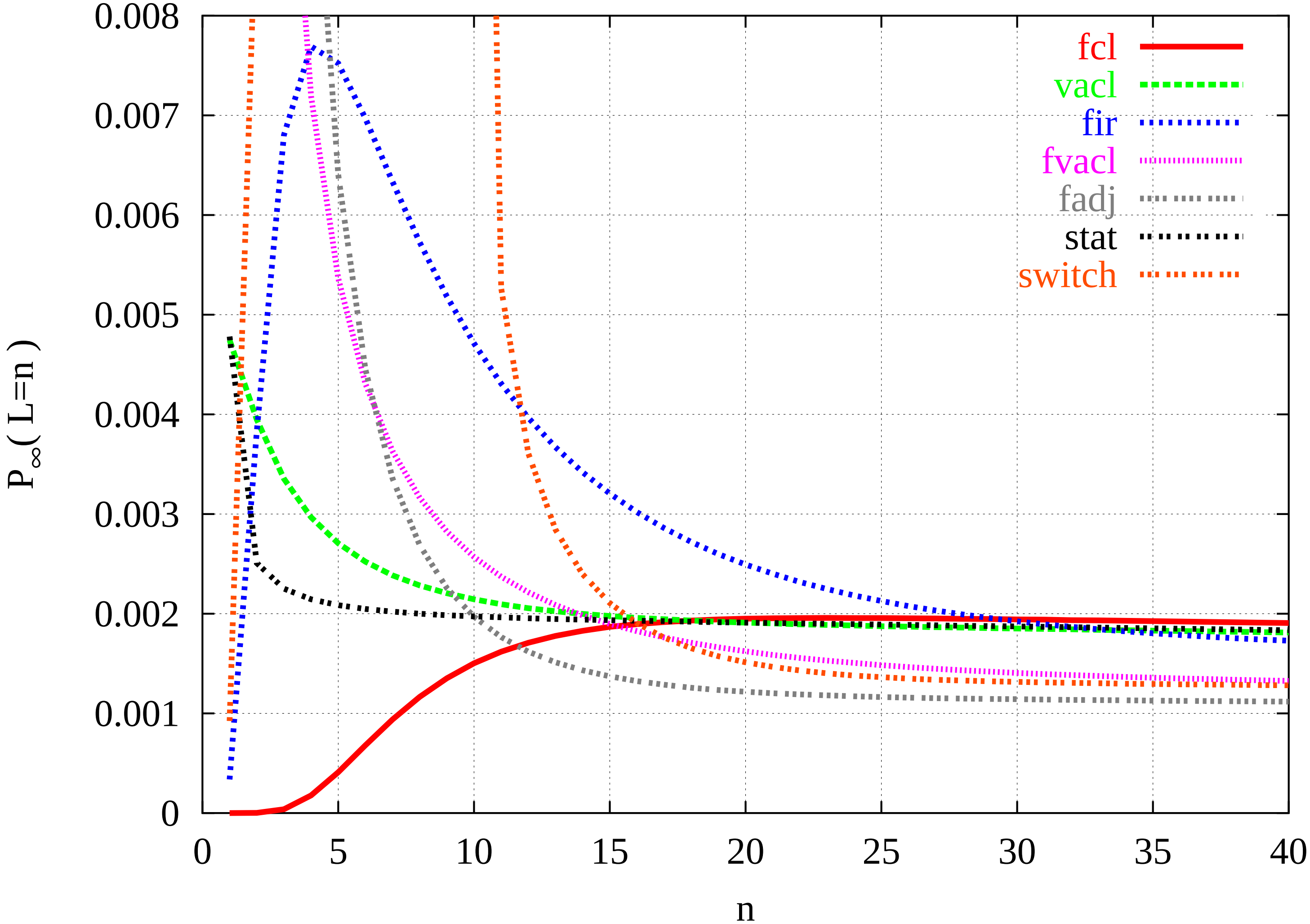
$$D_m = E_m(L - m + 1 | L \geq m)$$

$$(\lambda = 0.1,\ E_\infty(L) = 500,\ \delta = 1)$$



unter-Kontrolle-Wahrscheinlichkeitsfunktion von L

$(\lambda = 0.1,\ E_\infty(L) = 500)$



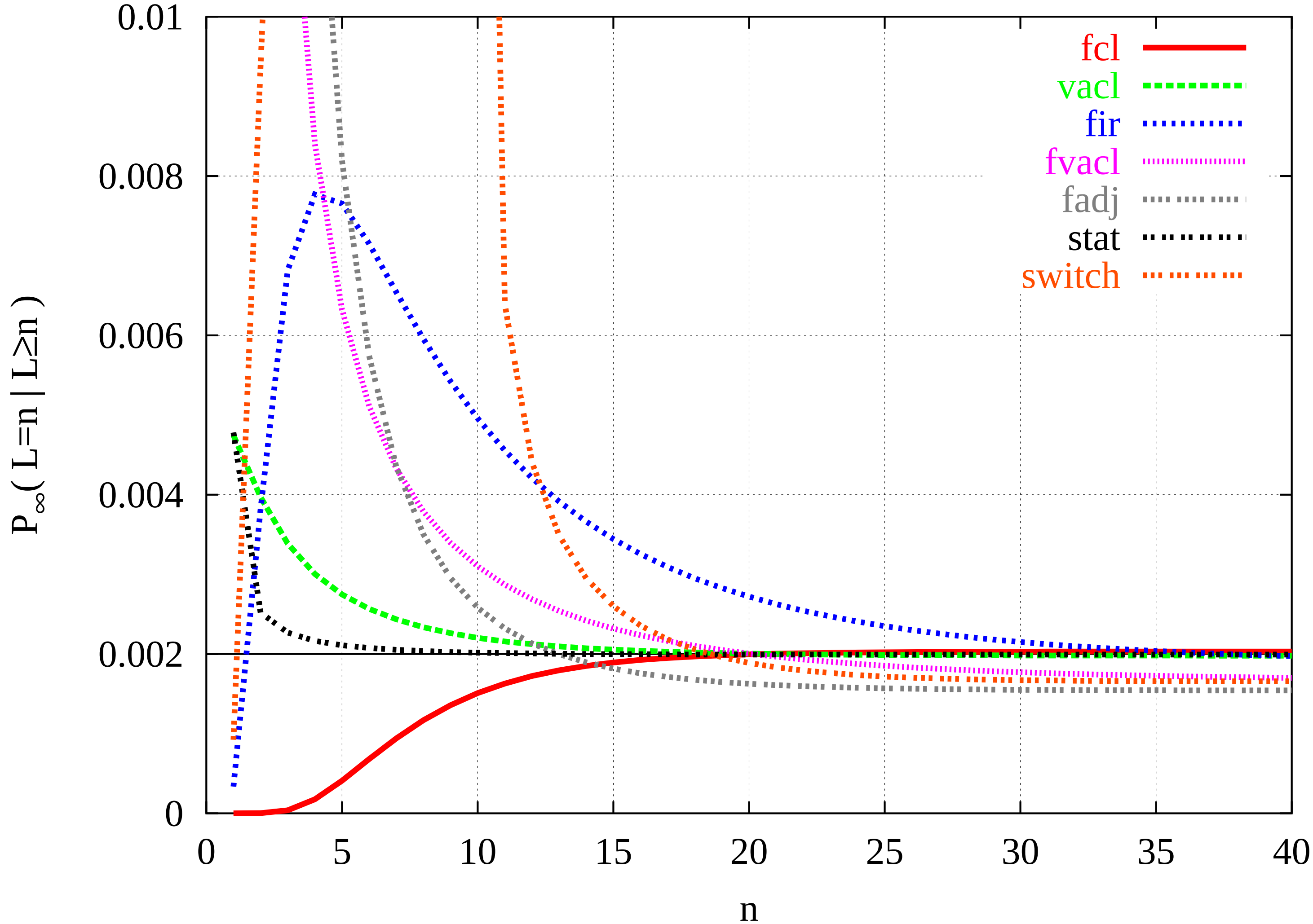
Wahrscheinlichkeiten für zeitige blinde Alarme

$\lambda = 0.1$, unter-Kontrolle-ARL, d. h. $E_\infty(L) = 500$

EWMA-Karte	fcl	vacl	fir	fvacl	fadj	stat	switch
$P_\infty(L=1)$	0.0000	0.0047	0.0003	**0.1125**	**0.1452**	0.0048	0.0009
$P_\infty(L=2)$	0.0000	0.0040	0.0038	0.0217	0.0435	0.0025	0.0094
$P_\infty(L=3)$	0.0000	0.0034	0.0068	0.0109	0.0190	0.0022	0.0170
$P_\infty(L \le 10)$	0.0063	0.0293	0.0551	**0.1742**	**0.2391**	0.0238	**0.1761**
				(0.02)			

Blinder-Alarm-Rate $r_n = P_\infty(L = n | L \geq n)$

$(\lambda = 0.1,\ E_\infty(L) = 500)$



Zusammenfassung

Bzgl. *fast initial response (FIR)* kann man drei Typen von EWMA-Kontrollkarten unterscheiden:

Eigenschaft	**"echtes" FIR**	**"moderates" FIR**	**balanciert**
anfängliche Detektionsempfindlichkeit	hoch	vergrößert	niedrig
Langzeitverhalten	wenig gestört	etwas gestört	gut
anfängliche Blinder-Alarm-Rate	hoch	moderat	niedrig
	fvacl `(Steiner)`	fir `(Lucas/Sacc.)`	stat `(Knoth)`
	switch `(Fink)`		fcl
	fvacl `(Rhoads et al.)`		vacl