19th International Conference on Computational Statistics

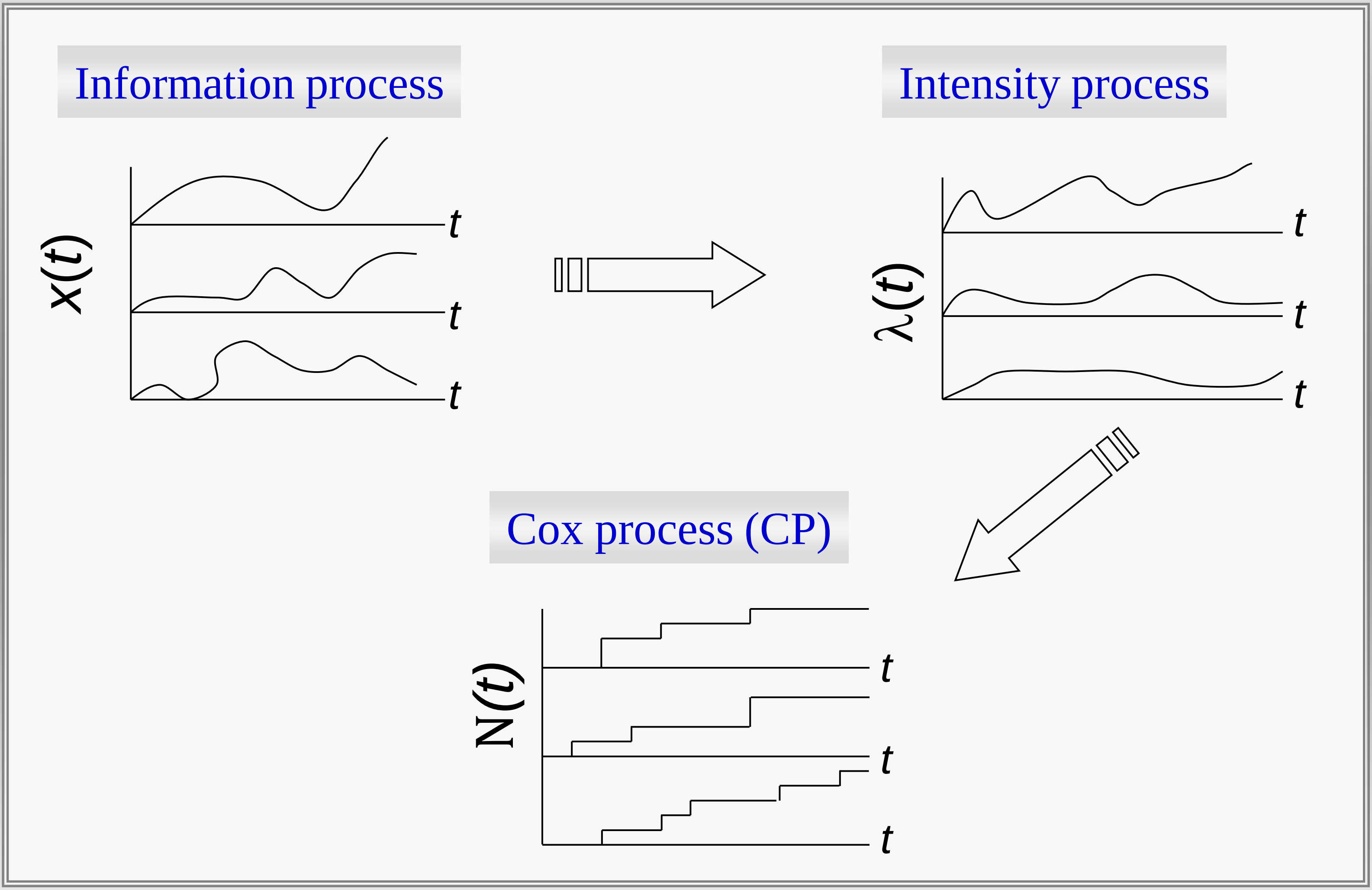
COMPSTAT 2010

Forecasting a Compound Cox Process by means of PCP

Paula R. Bouzas, Juan E. Ruiz-Castro
University of Granada, Spain

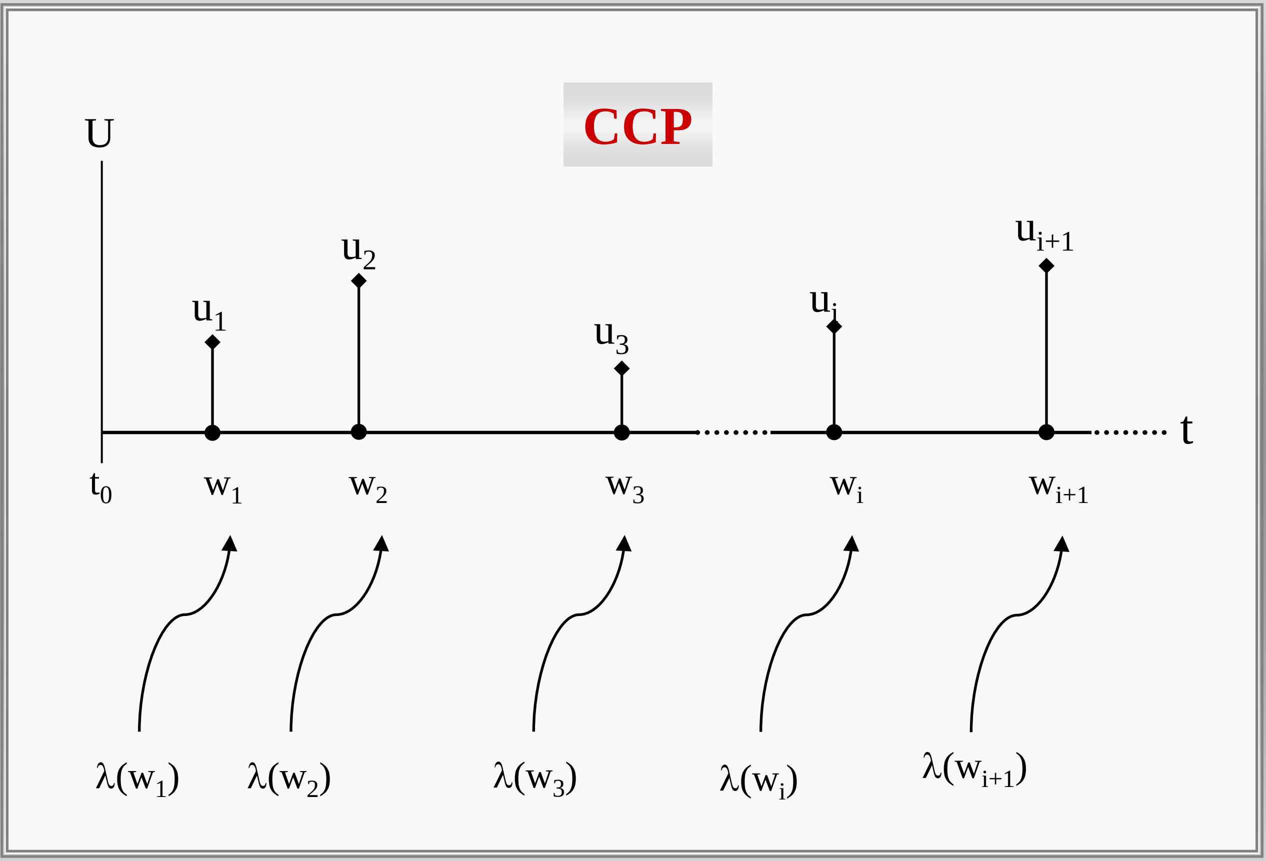
Nuria Ruiz-Fuentes
University of Jaén, Spain

Cox Process ≡ $N(t)$



Notation: $\lambda(t, x(t)) \equiv \lambda(t)$

Compound Cox Process $\equiv N(t)$



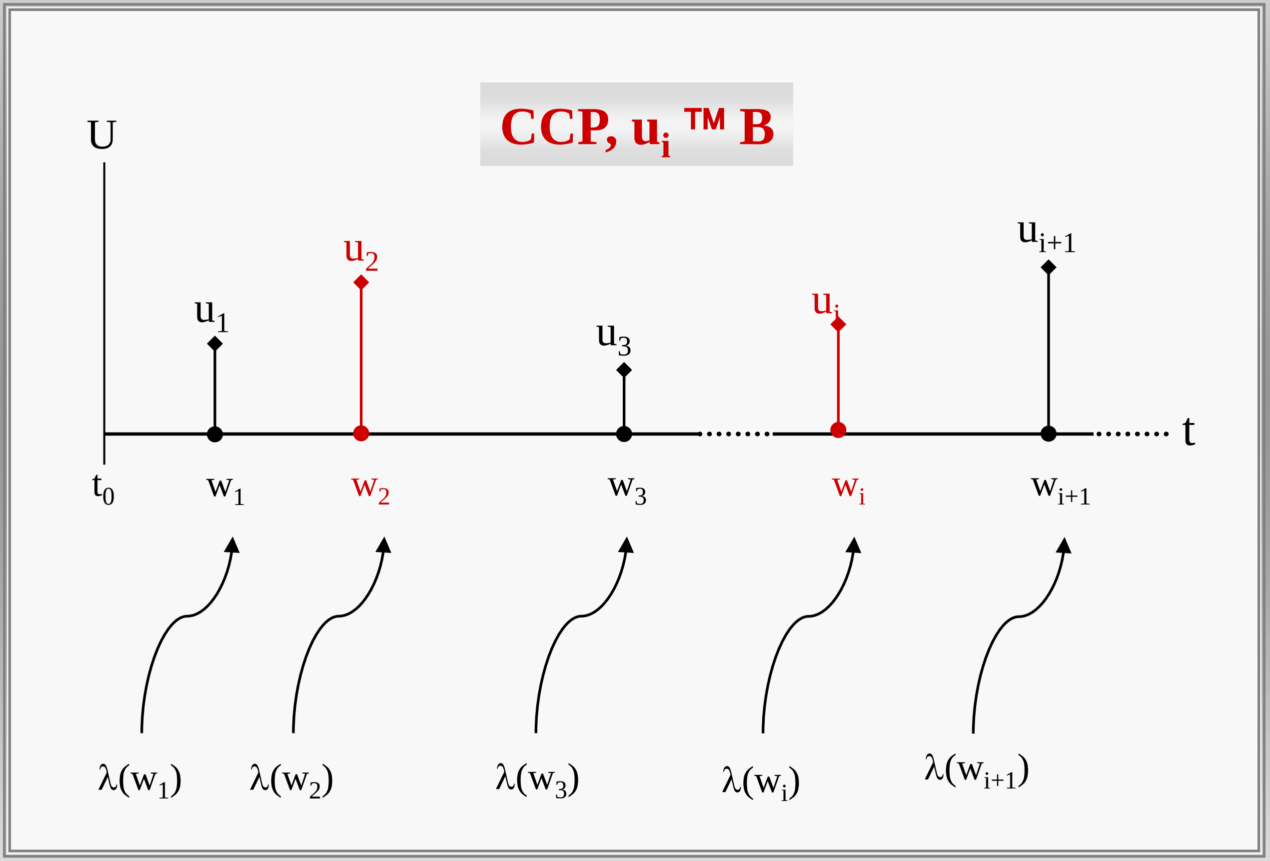
w_i = i-th ocurrence time

U r.v. in Y = mark space

$\lambda(w_i)$ = intensity in time w_i

u_i = mark of the i-th ocurrence

Compound Cox Process with marks in a given subset $\equiv$ $N(t,B)$



w_i = i-th ocurrence time

B ™ U r.v. in Y = mark space

$\lambda(w_i)$ = intensity in time w_i

u_i = mark of the i-th ocurrence

Representation theorems of a CCP
(Bouzas et al., 2007)

$N(t)$ is a CCP with intensity $\lambda(t)$ or mean $\Lambda(t)$

⇓

$N(t,B)$ is a CP with intensity $\lambda(t)\int_B P_u(dU)$

or mean $\Lambda(t)\int_B P_u(dU)$

Examples:

- Earthquakes of a certain magnitude interval or in a certain zone, ...
- Number of telephone calls with length in a given range
- Number of maximum prices of a stock beyond a given threshold
- Etc.

Counting statistics

- **Probability mass function**

$$P\left[N(t,B)=n\right]=E\left\{\frac{1}{n!}\left[\Lambda(t)\int_B P_u(dU)\right]^n \exp\left[-\Lambda(t)\int_B P_u(dU)\right]\right\};\ n=0,1,2,\ldots$$

- **Mean**

$$E\left[N(t,B)\right]=E\left[\Lambda(t)\int_B P_u(dU)\right]$$

- **Mode**

$$n_{\max}(t)=\max\left\{P\left[N(t,B)=n\right]\right\}=\begin{cases}\Lambda(t)\int_B P_u(dU)-1, & \Lambda(t)\int_B P_u(dU)\in\mathbb{N}\\ \left\{\begin{array}{l}\mathrm{int}\left[\Lambda(t)\int_B P_u(dU)-1\right],\\ or\\ \mathrm{int}\left[\Lambda(t)\int_B P_u(dU)-1\right]+1,\end{array}\right. & \Lambda(t)\int_B P_u(dU)\notin\mathbb{N}\\ 0, & \Lambda(t)\int_B P_u(dU)<1\end{cases}$$

Estimation of the mean process of *N*(t,B)

Estimation of the mean process of a CP by an *ad hoc* FPCA
(Bouzas et al., 2006)

$$\Lambda^q(t) = \mu_\Lambda(t)\sum_{j=1}^{q}\xi_j f_j(t),\ t \in I$$

Representation theorems of a CCP
(Bouzas et al., 2007)

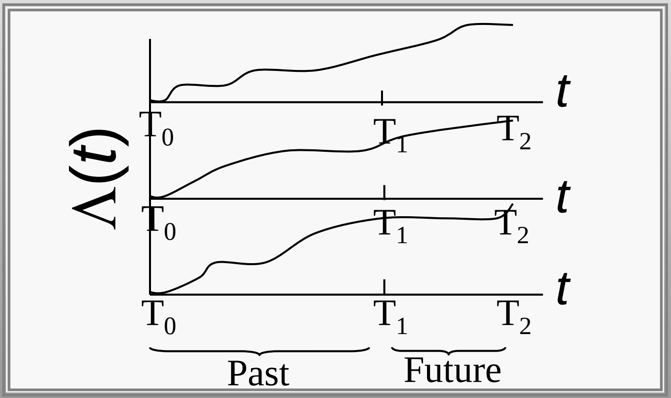
$N(t,B)$ is a CP with mean

$$\Lambda(t)\int_B P_u(dU)$$

$$\widehat{\Lambda(t)\int_B P_u(dU)} = \Lambda^q(t)\int_B P_u(dU)$$

$$t \in I$$

Forecasting the mean process of N(t,B)



Principal Components Prediction

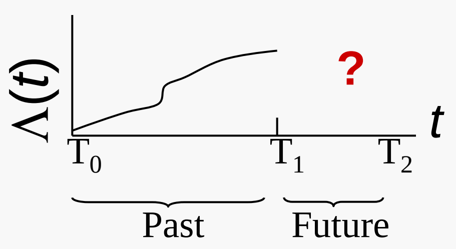
$$\Lambda^{q_1}(t) = \mu_\Lambda^1(t)\sum_{j=1}^{q_1} \xi_j f_j(t);\ t \in [T_0, T_1)$$

$$\Lambda^{q_2}(s) = \mu_\Lambda^2(s)\sum_{j=1}^{q_2} \eta_j g_j(s);\ \mathrm{s} \in [T_1, T_2)$$

$$\tilde{\Lambda}^{q_2}(s) = \mu_\Lambda^2(s)\sum_{j=1}^{q_2}\left(\sum_{i=1}^{p_j} b_i^j \xi_i\right) g_j(s);\ \mathrm{s} \in (T_1, T_2)$$

Forecasting the counting statistics of N(t,B)

$$E\left[N(s,B)\right] = E\left[\tilde{\Lambda}^{q_2}(s)\int_B P_u(dU)\right]$$



$$n_{\max}(s) = \begin{cases} \tilde{\Lambda}^{q_2}(s)\int_B P_u(dU) - 1, & \tilde{\Lambda}^{q_2}(s)\int_B P_u(dU) \in \mathbb{N} \\ \begin{cases} \text{int}\left[\tilde{\Lambda}^{q_2}(s)\int_B P_u(dU) - 1\right], \\ or \\ \text{int}\left[\tilde{\Lambda}^{q_2}(s)\int_B P_u(dU) - 1\right] + 1, \end{cases} & \tilde{\Lambda}^{q_2}(s)\int_B P_u(dU) \notin \mathbb{N} \\ 0, & \tilde{\Lambda}^{q_2}(s)\int_B P_u(dU) < 1 \end{cases}$$

Simulations

100 + 1 sample paths in [0,10]

1 $\lambda(t) \sim \Gamma(5,0.4)$; $U \sim B(10,0.4)$ and $\mathbf{B} = \{u;\ 4 \le u \le 6\}$

$$\Rightarrow \int_B P_u(dU) = p = 0.5630$$

$T_1 = 5,\ s = 7$

$$PCP(4;2,1,2,1)$$

$$E_s = 7.88 \qquad \hat{E}_s = 8.87$$

$$n_{\max} = 7 \qquad \hat{n}_{\max} = 8$$

$T_1 = 7,\ s = 8$

$$PCP(2;3,1)$$

$$E_s = 9.02 \qquad \hat{E}_s = 9.8$$

$$n_{\max} = 9 \qquad \hat{n}_{\max} = 8$$

Notation: Real mean $= E_s$, Estimated mean $= \hat{E}_s$,
Real mode $= n_{\max}$, Estimated mode $= \hat{n}_{\max}$

Simulations

100 + 1 sample paths in [0,10]

2

Boolean vector of four CP with λ(t) ~ Y(0,1);

U ~*lgn*(1,0.5) and **B** = {u; 2 ≤ u ≤ 5} $\Rightarrow \int_B P_u(dU) = p = 0.6188$

$T_1 = 4$, s = 5	$T_1 = 5$, s = 9
$PCP(3;2,2,1)$	$PCP(2;3,2)$
$E_s = 3.11 \quad \hat{E}_s = 3.93$	$E_s = 5.64 \quad \hat{E}_s = 6.33$
$n_{\max} = 3 \quad \hat{n}_{\max} = 1,3,4$	$n_{\max} = 5 \quad \hat{n}_{\max} = 4,7$

Notation: Real mean $= E_s$, Estimated mean $= \hat{E}_s$,

Real mode $= n_{\max}$, Estimated mode $= \hat{n}_{\max}$

Conclusions ≈ Basis

Representation theorems

CCP
with marks in a subset
N(t,B)

PCP models

•Prediction of the mean process

•Prediction of the mean

•Prediction of the mode

- Application to particular cases:
  CP with simultaneous ocurrences,
  multichannel CP, time-space CP,…
- Application to real data:
  turning points of a stock price,
  earthquakes with restrictives characteristics, …

References

◈ **Aguilera et al., 1997**. An aproximated principal component prediction model for continuous-time stochastic processes. *Appl. Stoch. Model. Data. Anal.*, 13, 61 – 72.

◈ **Aguilera et al., 1999**. Forecasting time series by functional PCA. Discussion of several weighted approaches. *Comput. Stat.*, 14, 443 – 467.

◈ **Bouzas et al., 2006**. Modelling the mean of a doubly stochastic Poisson process by functional data analysis. *Comput. Stat. Data Anal.*, 50, 2655 – 2667.

◈ **Bouzas et al., 2007**. Functional approach to the random mean of a compound Cox process. *Comput. Stat.*, 22, 467 – 479.

◈ **Ocaña et al., 1999**. Functional principal components analysis by choice of norm. *J. Multiv. Analysis*, 71, 262 – 276.