# Lecture Notes in Statistics

Edited by D. Brillinger, S. Fienberg, J. Gani, J. Hartigan, and K. Krickeberg

12

Martin Jacobsen

# Statistical Analysis of Counting Processes

Springer-Verlag
New York Heidelberg Berlin

# Lecture Notes in Statistics

Edited by D. Brillinger, S. Fienberg, J. Gani,
J. Hartigan, and K. Krickeberg

12

Martin Jacobsen

# Statistical Analysis of Counting Processes

Springer-Verlag
New York Heidelberg Berlin

Martin Jacobsen
University of Copenhagen
Institute of Mathematical Statistics
5 Universitetsparken
DK-2100 Copenhagen Ø
Denmark

AMS Classification: 62L99, 62M99

Library of Congress Cataloging in Publication Data

Jacobsen, Martin, 1942-
Statistical analysis of counting processes.

(Lecture notes in statistics ; 12)
Bibliography: p.
Includes index.
1. Stochastic processes. I. Title. II. Series:
Lecture notes in statistics (Springer-Verlag) ; v. 12.
QA274.J33 1982 519.5 82-19241



Printed and bound by R. R. Donnelley & Sons, Harrisonburg, VA.
Printed in the United States of America.

9 8 7 6 5 4 3 2 1

ISBN 0-387-90769-6 Springer-Verlag New York Heidelberg Berlin
ISBN 3-540-90769-6 Springer-Verlag Berlin Heidelberg New York

## Preface.

A first version of these lecture notes was prepared for a course given in 1980 at the University of Copenhagen to a class of graduate students in mathematical statistics. A thorough revision has led to the result presented here.

The main topic of the notes is the theory of multiplicative intensity models for counting processes, first introduced by Odd Aalen in his Ph.D. thesis from Berkeley 1975, and in a subsequent fundamental paper in the Annals of Statistics 1978.

In Copenhagen the interest in statistics on counting processes was sparked by a visit by Odd Aalen in 1976. At present the activities here are centered around Niels Keiding and his group at the Statistical Research Unit.

The Aalen theory is a fine example of how advanced probability theory may be used to develop a powerful, and for applications very relevant, statistical technique.

Aalen's work relies quite heavily on the 'theorie generale des processus' developed primarily by the French school of probability theory. But the general theory aims at much more general and profound results, than what is required to deal with objects of such a relatively simple structure as counting processes on the line. Since also this process theory is virtually inaccessible to non-probabilists, it would appear useful to have an account of what Aalen has done, that includes exactly the amount of probability required to deal satisfactorily and rigorously with statistical models for counting processes.

It has therefore been my aim to present a unified and essentially selfcontained exposition of the probability theory for counting processes and its application to the statistical theory of multiplicative intensity models. The inclusion of a purely probabilistic part conforms with my view that to apply the Aalen models in practice, one must have a thorough grasp of the underlying probability theory. Of course to carry out this programme, some knowledge of probability must be presupposed, especially conditional probabilities, weak convergence and basic martingale theory.

The first three chapters deal with univariate and multivariate counting processes and their probabilistic structure, while Chapters 4 and 5 are concerned with the definition of Aalen models and Aalen estimators, and the asymptotic results required to make the models applicable in statistical practice.

Naturally, the terminology and notation used in the general theory of processes has been carried over to the special situation treated here. One particularly relevant part of the general theory concerns the definition and basic properties of stochastic integrals of predictable processes with respect to martingales. This in particular, is one place where the setup involving only counting processes permits simplification compared to the general theory: whereas quite a lot of work is required to define the general stochastic integrals, all the integrals appearing here are ordinary (random) Lebesgue-Stieltjes integrals.

A number of exercises are given at the end of each chapter. Some of the exercises deal with proofs and arguments omitted from the text, while others aim at covering part of the theory and examples not included elsewhere.

Notation. The notation $s \downarrow\downarrow t$ means that $s \to t$ with $s > t$, where $s \downarrow t$ allows for $s \to t$ with $s \geq t$. For $X$ a random variable defined on a probability space $(\Omega, A, \mathbb{P})$, the notation $\mathbb{P}X$ rather than $EX$ is used for the expectation of $X$. Also, $\mathbb{P}(X;A)$ denotes the integral $\int_A X d\mathbb{P}$. Throughout $\mathbb{P}$ refers to a probability on some abstract probability space, while the letter $P$ is reserved for probabilities on some specific spaces. The notation $F$, $F_t$ for $\sigma$-algebras and $N_t$ for random variables also refers exclusively to these particular spaces.

Acknowledgements. I am especially indebted to Niels Keiding whose informal notes for a course on counting processes he gave in 1977-78 have been instrumental for the preparation of my own course in 1980, and thereby also for the writing of these notes.

I would like to thank Per Kragh Andersen, Richard Gill, Inge Henningsen, Søren Johansen, Niels Keiding, Henrik Ramlau-Hansen as well as my class, autumn 1980, for helpful discussions and comments.

The manuscript was typed at the Department of Mathematics, University of Copenhagen. I am most grateful for this essential assistance, and I am happy to thank Dita Andersen and Jannie Larsen for the very efficient job they have done.

Finally, thanks also to Simon Holmgaard for proofreading the entire manuscript.

Copenhagen, March 1982

Martin Jacobsen

TABLE OF CONTENTS

# 1. ONE-DIMENSIONAL COUNTING PROCESSES

## 1.1. Probabilities on $(0,\infty]$.

Consider the half-line $(0,\infty]$ (0 excluded, $\infty$ included) equipped with the Borel $\sigma$-algebra $\mathcal{B}$ of subsets generated by the subintervals of $(0,\infty]$.

A probability $Pr$ on $((0,\infty],\mathcal{B})$ may be described by its distribution function $F$, defined by $F(t) = Pr(0,t]$, $0 < t < \infty$. The function $F$ is non-decreasing, right-continuous and satisfies $F \leq 1$, $\lim_{t\downarrow\downarrow 0} F(t) = 0$. If conversely $F$ defined on $(0,\infty)$ is any function with these properties, then there is a unique probability $Pr$ on $(0,\infty]$ with $F$ as distribution function.

Instead of the distribution function $F$, one may use the survivor function $G = 1-F$. The following properties characterizes the survivor functions $G$ for probabilities on $(0,\infty]$: $G$ is non-increasing, right-continuous and satisfies $G \geq 0$, $\lim_{t\downarrow\downarrow 0} G(t) = 1$.

The termination point $t^\dagger$ of a probability on $(0,\infty]$ is defined as $t^\dagger = \inf\{t > 0: G(t) = 0\} = \sup\{t > 0: G(t) > 0\}$. Thus, if $t^\dagger < \infty$, $G(t^\dagger) = 0$ while $G(s) > 0$ for $s < t^\dagger$.

A probability on $(0,\infty]$ allows absorption if it has an atom at $\infty$: $Pr\{\infty\} > 0$. In that case $Pr\{\infty\} = G(\infty -) \stackrel{D}{=} \lim_{t\uparrow\uparrow\infty} G(t)$ is the absorption probability.

Suppose now that the probability $Pr$ on $(0,\infty]$ is absolutely continuous (strictly speaking the restriction of $Pr$ to $(0,\infty)$ is absolutely continuous with respect to Lebesgue measure) with density $f$, i.e. there is a non-negative, possibly infinite, measurable function $f$ defined on $(0,\infty)$ such that

$$F(t) = \int_0^t ds\ f(s) \qquad (0 < t < \infty)$$

(equivalently, $G(t) = G(\infty -) + \int_t^\infty ds\ f(s)$ for $0 < t < \infty$).

We shall say that Pr has a <u>smooth density</u> if $f: (0,\infty) \to [0,\infty]$ may be chosen to be right-continuous with left-limits everywhere such that $\lim_{t\downarrow\downarrow 0} f(t)$ exists (using the usual topology on $(0,\infty)$ and on $[0,\infty]$ the topology obtained when adjoining $\infty$ to $[0,\infty)$ (with the usual topology) in a one-point compactification).

Suppose Pr has a smooth density $f$. The <u>intensity</u> or <u>hazard</u> for Pr is the function $\mu: (0,\infty) \to [0\ \infty]$ defined by

$$\mu(t) = \begin{cases} f(t)/G(t) & \text{if } G(t) > 0 \\ 0 & \text{if } G(t) = 0 . \end{cases}$$

Since $f$ is right-continuous one has, provided $G(t) > 0$, that

$$\mu(t) = \lim_{h\downarrow\downarrow 0} \frac{1}{h} \Pr(t,t+h]/\Pr(t,\infty]$$

so that, suitably normalized, the intensity $\mu(t)$ measures the risk of "dying immediately after time $t$ given survival up to $t$".

1.1. <u>Example</u>. Let $0 \le \mu < \infty$ be a constant. The <u>exponential law</u> with rate $\mu$ is the probability on $(0,\infty]$ with survivor function $G(t) = \exp(-\mu t)$. It has smooth density $f(t) = \mu e^{-\mu t}$ and an intensity which is constant and equal to $\mu$. The special case $\mu = 0$ corresponds to the probability degenerate at $\infty$, (absorption probability 1). ∎

Expressed in terms of the survivor function $G$ alone, it is seen that

$$\mu = D^{+}(-\log G),$$

where $D^{+}$ is the right sided differential operator: $D^{+}\phi(t) = \lim_{h\downarrow\downarrow 0} \frac{1}{h}(\phi(t+h) - \phi(t))$. Conversely $G$ may be recovered from $\mu$ by

$$\text{(1.2)} \qquad G(t) = \exp\left(-\int_0^t ds\ \mu(s)\right) \qquad (0 < t < \infty).$$

It should now be clear that the intensity function $\mu$ for probabilities on $(0,\infty]$ with smooth densities are characterized by the following properties: $\mu$ is non-negative, right-continuous everywhere with left-limits everywhere except possibly at $t^\dagger$, the limit $\lim_{t\downarrow\downarrow 0}\mu(t)$ exists, $\mu$ is locally integrable at $0$ in the sense that $\int_0^h ds\mu(s) < \infty$ for some $h > 0$, and finally $\mu(t) = 0$ whenever $\int_0^t ds\mu(s) = \infty$.

If Pr has intensity $\mu$ it is seen that 1): Pr has a finite termination point $t^\dagger$ iff $\mu$ is not locally integrable, i.e. $\int_0^t ds\,\mu(s) = \infty$ for some $0 < t < \infty$, and in that case $\int_0^{t^\dagger} ds\,\mu(s) = \infty$ and $\int_0^t ds\,\mu(s) < \infty$ for $t < t^\dagger$; 2): Pr has $\infty$ as termination point but does not allow absorption iff $\mu$ is locally but not globally integrable, i.e. $\int_0^t ds\,\mu(s) < \infty$ for $0 < t < \infty$ and $\int_0^\infty ds\,\mu(s) = \infty$; 3): Pr allows absorption iff $\mu$ is globally integrable, i.e. $\int_0^\infty ds\,\mu(s) < \infty$, and in that case the absorption probability equals $\exp\left(-\int_0^\infty ds\,\mu(s)\right)$.

If for some $t_0 > 0$, $Pr(t_0,\infty] = 1$, then of course $\mu(t) = 0$ for $0 \le t < t_0$ and (1.2) may be written

$$G(t) = \exp\left(-\int_{t_0}^{t} ds\,\mu(s)\right) \qquad (t_0 \le t < \infty)$$

with $G(t) = 1$ for $t \le t_0$.

1.3. Example. If Pr has intensity $\mu$, then for any $t_0 > 0$ the conditional probability $Pr(\cdot\,|(t_0,\infty])$ has intensity function

$$\mu_{|t_0}(t) = \begin{cases} 0 & (0 < t < t_0) \\ \mu(t) & (t_0 \le t < \infty) \end{cases}$$

and survivor function

$$G_{|t_0}(t) = \begin{cases} 1 & (0 < t < t_0) \\ \exp(-\int_{t_0}^{t} ds\,\mu(s)) & (t_0 \le t < \infty). \end{cases} \qquad \square$$

The following result will be useful later.

1.4. Proposition. Let $T$ be a $(0,\infty]$-valued random variable such that the distribution of $T$ has a smooth density with intensity $\mu$ and let $0 < \mu_0 < \infty$ be a constant. Then, assuming that $\int_0^\infty ds\ \mu(s) = \infty$, the random variable

$$U = \frac{1}{\mu_0} \int_0^T ds\ \mu(s)$$

follows an exponential law with rate $\mu_0$.

Proof. Define $H(t) = \int_0^t ds\ \mu(s)$ and denote by $H^{-1}$ the right-continuous inverse of $H$: $H^{-1}(u) = \inf\{t>0: H(t) > u\}$. Since $\int_0^\infty ds\ \mu(s) = \infty$, $H^{-1}(u)$ is defined for all $0 \leq u < \infty$ and furthermore satisfies $H(H^{-1}(u)) = u$, $H(t) > u$ for $t > H^{-1}(u)$. Thus, if $\mathbb{P}$ denotes the probability on the probability space where $T$ is defined, for any $0 \leq u < \infty$

$$\begin{aligned} \mathbb{P}(U > u) &= \mathbb{P}(H(T) > \mu_0 u) = \mathbb{P}(T > H^{-1}(\mu_0 u)) \\ &= \exp(-H(H^{-1}(\mu_0 u))) = e^{-\mu_0 u}. \end{aligned}$$

∎

## 1.2. The definition of one-dimensional counting processes.

A one-dimensional counting process may be thought of as a stochastic process recording at any given time $t$ the number of certain events having occurred before time $t$. This is formalized in Definition 2.1 below.

Let $(\Omega, A, A_t, \mathbb{P})$ be a probability space with a filtration, i.e. $(\Omega, A, \mathbb{P})$ in a usual probability space and $(A_t)_{t\geq 0}$ is a family of sub $\sigma$-algebras of $A$ such that $A_s \subset A_t$ when $s \leq t$. A stochastic process $X = (X_t)_{t\geq 0}$ defined on $(\Omega, A)$ is adapted to $(A_t)$ if each $X_t$ is $A_t$-measurable.

(Note: when writing $(I_t)_{t\geq 0}$ for some indexed family of objects, the indexing set is $[0,\infty)$, so there is an $I_t$ for each $0 \leq t < \infty$ but not apriori for $t = \infty$).

2.1. Definition. A one-dimensional counting process on a filtered probability space $(\Omega, A, A_t, \mathbb{P})$, is an adapted stochastic process $K = (K_t)_{t\geq 0}$, each $K_t$ taking values in $\overline{\mathbb{N}}_0 = \{0,1,\ldots,\infty\}$ with $\mathbb{P}(K_0 = 0) = 1$ and such that almost all sample paths are non-decreasing and right-continuous everywhere, increasing only by jumps of size 1.

The process is stable if $\mathbb{P}(K_t < \infty) = 1$ for all $t \geq 0$.

The process allows absorption if $\mathbb{P}(\sup_{t\geq 0} K_t < \infty) > 0$. ∎

Recall that the sample paths for $K = (K_t)_{t\geq 0}$ are the functions $t \to K_t(\omega)$ obtained for any $\omega \in \Omega$. The definition demands that for $\omega$ outside a $\mathbb{P}$-null set, the sample path determined by $\omega$ be right-continuous. The topology on $\overline{\mathbb{N}}_0$ to be referred to when making this statement precise is the one obtained by adjoining $\infty$ as the one-point compactification to $\mathbb{N}_0 = \{0,1,\ldots\}$, the set of non-negative integers, equipped with the discrete topology.

It is readily checked that with this choice for the topology on $\bar{\mathbb{N}}_0$, almost all sample paths will have left-limits everywhere.

Since we shall only discuss one-dimensional counting processes in this section we shall for simplicity refer to such a process as a counting process.

If we are just given a probality space $(\Omega, A, \mathbb{P})$ and a process $K = (K_t)_{t \geq 0}$ with almost all sample paths having the analytic properties required by Definition 2.1, it is always possible to find a filtration $(\mathcal{K}_t)_{t \geq 0}$ such that on $(\Omega, A, \mathcal{K}_t, \mathbb{P})$ is a counting process: define $\mathcal{K}_t = \sigma(K_s)_{s \leq t}$, the smallest sub $\sigma$-algebra of $A$ with respect to which all $K_s$, $s \leq t$ become measurable. If we are given a counting process $K$ on a filtered space $(\Omega, A, A_t, \mathbb{P})$, then $K$ is also a counting process on $(\Omega, A, \mathcal{K}_t, \mathbb{P})$ and $\mathcal{K}_t \subset A_t$ so that $(\mathcal{K}_t)_{t \geq 0}$ is the smallest filtration with respect to which $K$ is a counting process. We shall call $(\mathcal{K}_t)_{t \geq 0}$ the <u>self-exciting</u> filtration for the process $K$.

Given a counting process $K = (K_t)_{t \geq 0}$, consider the mapping $\omega \to (K_t(\omega))_{t \geq 0}$ which to every $\omega \in \Omega$ associates the corresponding sample path of the process. This mapping $T$ carries each $\omega$ into an element of the function space $\bar{\mathbb{N}}_0^{[0,\infty)}$ of all functions (paths) defined on $[0,\infty)$ taking values in $\bar{\mathbb{N}}_0$, which, for almost all $\omega$, has specific analytic properties. Taking out a relevant subset $W$ of $\bar{\mathbb{N}}_0^{[0,\infty)}$ and equipping it as a measurable space one may therefore transform the original probability $\mathbb{P}$ on $\Omega$ into a probability $P = T(\mathbb{P})$ on $W$, which in a canonical fashion describes the probabilitic properties of the process $K$. These considerations lead to Definitions 2.2 and 2.3.

2.2. <u>Definition</u>. The <u>full counting process path-space</u> is the subset $\bar{W}$ of $\bar{\mathbb{N}}_0^{[0,\infty)}$ consisting of those paths $w: [0,\infty) \to \bar{\mathbb{N}}_0$ with $w(0) = 0$ which are everywhere right-continuous and non-decreasing, increasing only in jumps of size 1.

The stable counting process path-space is the subset $W$ of $\overline{W}$ consisting of those paths $w \in \overline{W}$ for which $w(t) < \infty$ for all $t \geq 0$.

▮

From purely theoretical considerations, the full space $\overline{W}$ is the natural one to use as will be apparent from the next subsection. But for most statistical applications the stable space $W$ is the appropriate one.

For $t \geq 0$, define $N_t : \overline{W}\ (W) \to \overline{\mathbb{N}}_0$ by $N_t(w) = w(t)$ and let $F$ denote the smallest $\sigma$-algebra of subsets of $\overline{W}\ (W)$ that makes all $N_t$ measurable: $F = \sigma((N_t = n) : n \in \mathbb{N}_0,\ t \geq 0)$. Also, let $F_t$ be the $\sigma$-algebra generated by $(N_s)_{s \leq t}$. (Thus $(F_t)$ is a filtration on $\overline{W}\ (W)$ and $N = (N_t)$ is adapted to this filtration). Note that $F_0 = \{\emptyset, \overline{W}\}$ $(\{\emptyset, W\})$. Finally define $N_\infty = \lim_{t \uparrow \uparrow \infty} N_t$.

It is possible to describe the $\sigma$-algebra $F_t$ in a different way. On $\overline{W}\ (W)$ introduce an equivalence relation $\underset{t}{\sim}$ by requiring that $w \underset{t}{\sim} w'$ iff $w(s) = w'(s)$ for $0 \leq s \leq t$. Then $F \in F_t$ iff $F \in F$ and $F$ is a union of $\underset{t}{\sim}$-equivalence classes, which are then also referred to as the atoms of $F_t$. (*Sketch of proof: clearly the collection of sets which are $F$-measurable unions of $\underset{t}{\sim}$-equivalence classes, form a $\sigma$-algebra $\tilde{F}_t$, and since obviously $(N_s = n) \in \tilde{F}_t$ for $n \in \mathbb{N}_0$, $0 \leq s \leq t$, we have $F_t \subset \tilde{F}_t$. Conversely the mapping $S_t : \overline{W} \to \overline{W}$ given by $N_s \circ S_t = N_{s \wedge t}$ $(s \geq 0)$ is $F_t$-measurable and has the property that $S_t^{-1} F = F$ for $F \in \tilde{F}_t$ wherefore also $\tilde{F}_t \subset F_t$*). Using this equivalence class description of $F_t$, to show that a random variable defined on $(\overline{W}, F)$ $((W, F))$ is $F_t$-measurable, amounts to showing that it is constant on each $F_t$-atom. The $\sigma$-algebra $F_t$ contains all information about the behaviour of the $N_s$ on the time interval $[0,t]$. It is customary in general process theory to consider the slightly larger $\sigma$-algebras $F_{t+} \overset{D}{=} \bigcap_{\varepsilon > 0} F_{t+\varepsilon}$. However, in the case of the counting process path spaces $\overline{W}$ and $\overline{W}$, we have that $F_{t+} = F_t$, the reason being that knowing exactly the behaviour of a path $w$ on $[0,t]$ tell us also the behaviour of $w$ on $[0,t+\varepsilon]$ for some $\varepsilon > 0$, viz. $w(s) = w(t)$ for $t \leq s \leq t+\varepsilon$ by right-continuity. (Formally a proof that $F_{t+} = F_t$ may be given as follows: it is shown that $F_{t+}$ consists of the sets which are $F$-measurable unions of equivalence classes for the equivalence relation $\underset{t+}{\sim}$ given by $w \underset{t+}{\sim} w'$ iff for some $\varepsilon = \varepsilon(w,w') > 0$, $w \underset{t+\varepsilon}{\sim} w'$; then it is observed that $\underset{t+}{\sim}$ is the same as $\underset{t}{\sim}$).

We have now equipped the path-spaces $\overline{W}$ and $W$ with a measurable structure and are ready to give the next fundamental definition.

2.3. Definition. A canonical one-dimensional counting process is a probability on $(\overline{W},F)$. A stable canonical one-dimensional counting process is a probability on $(W,F)$. ∎

For convenience we shall abbreviate canonical counting process as CCP.

Thus, for CCP's the family of random variables defining the process is always the family $(N_t)$ of projections and a CCP is cahracterized exclusively as a probability on $\overline{W}$ or $W$.

If $P$ is a CCP we shall also use the symbol $P$ to denote $P$-expectation. Thus, if $F \in F$ and $U$ is real-valued and $F$-measurable we write $P(F)$, $P(U)$, $P(U;F)$ for respectively the $P$-measure of the set $F$, the integral $\int dP\, U$ and the integral $\int_F dP\, U$.

Note that any CCP, $P$, is completely determined by its collection of finite-dimensional distributions, i.e. the $P$-distribution of any vector $(N_{t_1},\ldots,N_{t_r})$ where $r \in \mathbb{N}$, $0 \leq t_1 < \ldots < t_r$.

Suppose that $K = (K_t)_{t\geq 0}$ is a counting process on $(\Omega, A, A_t, \mathbb{P})$ in the sense of Definition 2.1. Taking away a $\mathbb{P}$-null set $N$, the mapping $T$ discussed above becomes a measurable mapping from

$(\Omega\setminus N, A(\Omega\setminus N))$ to $(\overline{W},F)$ (to $(W,F)$ if $K$ is stable) and hence induces a probability $P = T(\mathbb{P})$ on $(\overline{W},F)$ $((W,F))$, the canonical counting process generated by $K$.

By the transformation some information may have been lost, but all information contained in the process itself has been retained: for every $t \geq 0$, knowing the restriction of $P$ to $F_t$ determines the restriction of $\mathbb{P}$ to $K_t$, and complete knowledge of $P$ determines the restriction of $\mathbb{P}$ to $K$, the smallest sub $\sigma$-algebra of $A$ containing the members $K_t$ of the self-exciting filtration.

In these notes we shall mainly be concerned with CCP's. In statistical terms this means that we shall consider only the counting process itself as observable.

2.4. Example. The most important of all counting processes is the Poisson process. For $0 < \mu < \infty$ a constant, the canonical Poisson process with rate (or intensity) $\mu$ is the probability $\Pi_\mu$ on the stable space $(W,F)$ with respect to which $(N_t)_{t\geq 0}$ has stationary independent Poisson increments: for $r \in \mathbb{N}$, $0 = t_0 < t_1 < \ldots < t_r$, $n_1,\ldots,n_r \in \mathbb{N}_0$

$$\Pi_\mu(N_{t_i}-N_{t_{i-1}} = n_i,\ i=1,\ldots,r) = \prod_{i=1}^{N} \Pi_\mu(N_{t_i}-N_{t_{i-1}} = n_i)$$

and for $0 \leq s < t$, $n \in \mathbb{N}_0$

$$\Pi_\mu(N_t-N_s = n) = \frac{(\mu(t-s))^n}{n!} e^{-\mu(t-s)} .$$

These distributional properties may also be written

$$\Pi_\mu(N_u-N_t = n\,|\,F_t) = \frac{(\mu(u-t))^n}{n!} e^{-\mu(u-t)}$$

for $0 \leq t \leq u$, $n \in \mathbb{N}_0$. ∎

A CCP $P$ is Markov if for all $t < u$, $n \in \mathbb{N}_0$