NASA CONTRACTOR
REPORT

NASA CR-129044

APPLICATION OF MARKOV CHAIN THEORY TO ASTP NATURAL ENVIRONMENT LAUNCH CRITERIA AT KENNEDY SPACE CENTER

By M. E. Graves and M. Perlmutter
Northrop Services, Inc.
P.O. Box 1484
Huntsville, Alabama 35807

September 1974

Prepared for

NASA-GEORGE C. MARSHALL SPACE FLIGHT CENTER
Marshall Space Flight Center, Alabama 35812

(NASA-CR-129044) APPLICATION OF MARKOV CHAIN THEORY TO ASTP NATURAL ENVIRONMENT LAUNCH CRITERIA AT KENNEDY SPACE CENTER (Northrop Services, Inc., Huntsville, Ala.) 29 p HC $3.25 CSCL 22D G3/30

N74-35248

Unclas
51080

TECHNICAL REPORT STANDARD TITLE PAGE

1. REPORT NO.	2. GOVERNMENT ACCESSION NO.	3. RECIPIENT'S CATALOG NO.
NASA CR-129044		

4. TITLE AND SUBTITLE
Application of Markov Chain Theory to ASTP Natural Environment Launch Criteria at Kennedy Space Center

5. REPORT DATE
September 1974

6. PERFORMING ORGANIZATION CODE

7. AUTHOR(S)
M. E. Graves and M. Perlmutter

8. PERFORMING ORGANIZATION REPORT #

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Northrop Services, Inc.
P. O. Box 1484
Huntsville, AL 35807

10. WORK UNIT NO.

11. CONTRACT OR GRANT NO.
NAS8-21810

12. SPONSORING AGENCY NAME AND ADDRESS
National Aeronautics and Space Administration
Washington, DC 20546

13. TYPE OF REPORT & PERIOD COVERED
Contractor Report

14. SPONSORING AGENCY CODE

15. SUPPLEMENTARY NOTES
Technical Coordinator: S. C. Brown, Aerospace Environment Division, Space Sciences Laboratory, Marshall Space Flight Center, Alabama, 35812.

16. ABSTRACT

To aid the planning of the Apollo Soyuz Test Program (ASTP), certain natural environment statistical relationships are presented, based on Markov theory and empirical counts. The practical results are in terms of conditional probability of favorable and unfavorable launch conditions at Kennedy Space Center (KSC). They are based upon 15 years of recorded weather data which are analyzed under a set of natural environmental launch constraints.

Three specific forecasting problems are treated, namely, (1) the length of record of past weather which is useful to a prediction; (2) the effect of persistence in runs of favorable and unfavorable conditions; (3) the forecasting of future weather in probabilistic terms.

The term "unfavorable conditions" applies to anytime during the afternoon rather than to a specific launch hour. This means the probability values are somewhat conservative — perhaps descriptive of marginal launch conditions.

The results indicate that unfavorable launch conditions occur on 30.6 percent of the July-August afternoons. If unfavorable conditions were reported this afternoon the chance of a repeat occurrence tomorrow afternoon increases to 51.7 percent.

Unfavorable and favorable launch conditions are best described by first and second order Markov chains, respectively.

17. KEY WORDS

18. DISTRIBUTION STATEMENT
Unclassified - Unlimited

C. A. Lundquist
Director
Space Sciences Laboratory

19. SECURITY CLASSIF. (of this report)	20. SECURITY CLASSIF. (of this page)	21. NO. OF PAGES	22. PRICE
Unclassified	Unclassified	28	NTIS

MSFC - Form 3292 (Rev December 1972)
For sale by National Technical Information Service, Springfield, Virginia 22151

FOREWORD

This document presents the results of work done by Northrop Services, Inc. under Contract NAS8-21810 to the Science and Engineering Directorate of the Marshall Space Flight Center. This work was conducted in response to the requirements of Appendix A, Schedule Order Number A02Z (A-14). Mr. S. C. Brown, ES42, is the NASA/MSFC Technical Coordinator for this task.

ACKNOWLEDGEMENT

The authors are grateful to Mr. S. Clark Brown, Dr. George H. Fichtl, and Mr. Orvel E. Smith of NASA for their guidance and fruitful discussion during the project.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

Section I
INTRODUCTION

The Apollo Soyuz Test Project (ASTP) currently schedules an Apollo launch at 1500 EST on 15 July 1975 from Kennedy Space Center. The launch vehicle is subject to certain constraints, and the deadline for a "go" or "no go" decision for July 15 (or any subsequent date created by mission postponement) is at the previous midnight, i.e., 15 hours earlier.

For mission planning, the effect of persistence in the weather events which cause operational delays is very important, so this study will take persistence into account in its presentation of conditional probabilities of launch conditions. Markov theory will be applied, and the results will be compared with empirical probabilities obtained directly from the data.

Other studies of persistence in meteorological events have been conducted by Feyerherm and Bark (ref. 1), Williams (ref. 2), Hopkins and Robillard (ref. 3), Weiss (ref. 4), Eriksson (ref. 5), Brelsford and Jones (ref. 6), Smith (ref. 7), Gabriel and Neumann (refs. 8, 9), and Caskey (ref. 10). Additional material on runs is available in Gabriel (ref. 11), Walker and Duncan (ref. 12), von Mises (ref. 13), Feller (ref. 14), and Wilks (ref. 15). A rather complete treatment of hypothesis-testing in regard to Markov chains is found in Anderson and Goodman (ref. 16).

As is well known, thunderstorms are a major impediment to launch operations at Kennedy Space Center in the summer. Two recent station studies of thunderstorms are the statistical investigations by Falls (ref. 17) and Neumann (ref. 18). The former paper reaches the conclusion that thunderstorm events at Cape Kennedy are well represented by a negative binomial distribution. The latter paper discusses methods for predicting thunderstorms at the Cape, and it presents prediction equations obtained by nonlinear regression.

The present study is divided into two main sections. Section II defines an ASTP unfavorable day and reveals the distribution of such "restrictions" over the 15-year data period which is available. Then it analyzes the utility of past weather conditions in advancing forward a single day with conditional probabilities. Section III applies Markov chain theory and in particular, the Chapman-Kolmogorov Equations, to the data, comparing the results with experimental counts. Both first and second order Markov processes are investigated, and a statistical method is presented which predicts launch conditions on a dichotomous basis for mission planning.

Finally, in Section IV a summary of the work is made and the conclusions are presented. The references are given in Section V.

Section II

ASTP WEATHER RESTRICTION

2.1 DEFINITION

An ASTP weather restriction is defined as the occurrence of any of the following conditions:

a. Precipitation

b. Thunderstorms with a cloud ceiling

c. Cumulus cloud ceiling < 4000 feet

d. Wind speed $\geq$ 25 knots at 30 feet

The record of hourly observations of weather elements at Cape Canaveral AFS from 1957 to 1971 is complete, and additional information is available on the peak wind speed each hour and the duration, location, and other characteristics of each thunderstorm. To treat just the 1500 EST observations would ignore a substantial number of cases when restrictive conditions occurred during the afternoon and/or the weather at 1500 EST was nearly restricted. Therefore, the writers have adopted a broader definition of "unfavorable conditions" such that one or more of the four constraints is simply required to be reported at least once in the hourly record from 1200 EST to 1700 EST, inclusively. This makes the concept of "unfavorable" almost synomymous with the term, "marginal".

A link can be provided between this definition and the probability of occurrence of restricted conditions at a particular hour. From all July and August data for the period 1957-1971, the relative frequency of restricted conditions at 1300 EST, 1400 EST, or 1500 EST, given that a restriction has taken place at least once between 1200 and 1700 EST, is 0.333, 0.390, or 0.434, respectively. The 95 percent confidence limits on each value are $\pm$ 0.049.

2.2 TIME DISTRIBUTION OF ASTP UNFAVORABLE DAYS

Five-day moving averages of the frequency of unfavorable days for an ASTP launch are calculated for each July and August day and shown in Figure

2-1. From the broken, visually-fitted curve, a rising trend is apparent in July with a reversal early in August. There are a few irregularities, but only the dip in mid-August is significant at the one percent level by Student's t-test. In view of these results, the data are considered to be sufficiently homogeneous to permit the inclusion of August weather reports. This produces a total of 930 days of observation for the 15-year period under study. A test for cycles in the occurrence of unfavorable days discloses that the data are free of such periodicities.

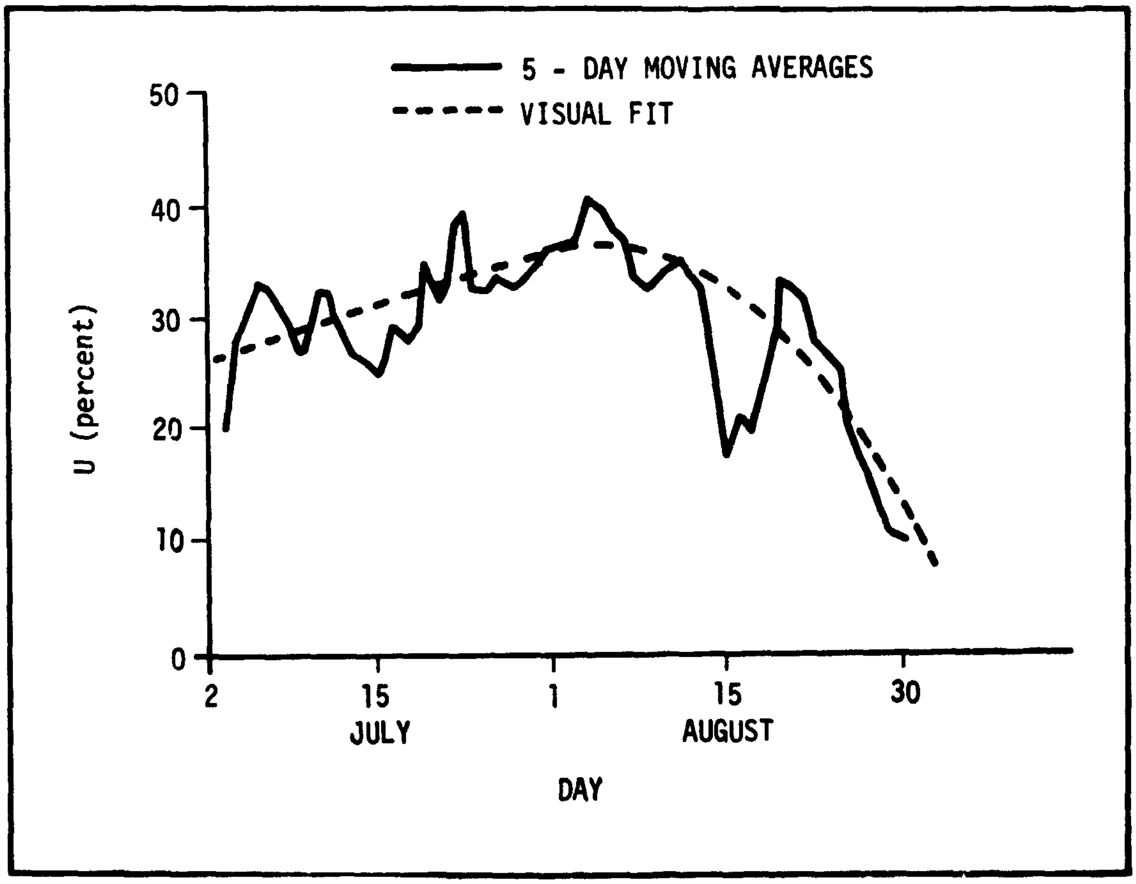


Figure 2-1. FREQUENCY OF DAYS WITH UNFAVORABLE CONDITIONS, KSC, 1957-1971

2.3 ANALYSIS OF A ONE-DAY ADVANCE

The following question will now be considered: If an advance of one day into the future is to be made in a prediction scheme based upon the past weather with its persistent nature, how many days of past weather can be profitably used?

2.3.1 Theoretical Aspects

Table 2-1 presents empirical and calculated values of favorable conditions for ASTP launch the following day, given present and previous days weather behavior. In the table $P(F_1)$ is the probability of the next day being favorable for the launch, independent of present or previous days weather. This is considered a zero order probability. $P(F_1|U_o)$ is the probability of the next day being favorable for launch, given that the weather today is unfavorable for launch. This is considered a first order probability. Similarly, $P(F_1|U_oF_{-1})$ is the probability that the next day will be favorable for launch given that today is unfavorable and the day before was favorable. This is considered a second order probability. The results of Table 2-1 are also plotted in Figure 2-2.

Table 2-1. EMPIRICAL PROBABILITIES AND CALCULATED PROBABILITIES OF FAVORABLE CONDITIONS FOR ASTP LAUNCH

	EMPIRICAL VALUE	CALCULATED VALUE	CALCULATED FROM
$P(F_1)$	0.694 ± 0.026*		
$P(U_1)$		0.306 ± 0.026*	P(F)
$P(F_1\|F_o)$	0.788 ± 0.028	0.787 ± 0.022*	$P(F_1\|U_o)$
$P(F_1\|U_o)$	0.483 ± 0.051*	0.481 ± 0.064	$P(F_1\|F_o)$
$P(F_1\|F_oF_{-1})$	0.823 ± 0.028	0.807 ± 0.020*	$P(F_1\|F_oU_{-1})$
$P(F_1\|F_oU_{-1})$	0.714 ± 0.070*	0.654 ± 0.103	$P(F_1\|F_oF_{-1})$
$P(F_1\|U_oU_{-1})$	0.493 ± 0.069*	0.492 ± 0.072	$P(F_1\|U_oF_{-1})$
$P(F_1\|U_oF_{-1})$	0.473 ± 0.077	0.472 ± 0.074*	$P(F_1\|U_oU_{-1})$

* *The subscripts indicate the order of days for a favorable (F) or unfavorable (U) case. The asterisks indicate the values used in subsequent calculations.*

The empirical probabilities are obtained by counting the number of days that fit the class description (r) divided by the total number of days in the data set (n). For example,

$$p(F_1 | U_o\ F_{-1}) \approx \frac{r}{n} \quad . \tag{1}$$

Thus by counting the number of days r in the total data set where a favorable launch day is followed by an unfavorable day and a preceding favorable day, and then dividing this by the total number of days in the data set, n, an approximation to $p(F_1 | U_o\ F_{-1})$ is obtained which improves as $n \to \infty$. In this limit $r/n \to p(F_1 | U_o\ F_{-1})$. Confidence limits are then calculated on p using Bayes Theorem as shown in Pratt, Raiffa and Schaefer (ref 19). There is some prior distribution of p assumed before the present data are utilized. This is called the prior distribution given by $f_I(p)$. Since all that is known is that p has a value between 0 and 1 and any value is equally likely, one can let

$$f_I(p) = 1 \tag{2}$$

The data, $D_{r,n}$, is then obtained. Then the posterior distribution of p, given the data $D_{r,n}$, $[f_F(p|D_{r,n})]$, can be obtained from Bayes Theorem

$$f_F(p|D_{r,n}) = \frac{f(D_{r,r}|p) f_I(p)}{f(D_{r,n})} \tag{3}$$

The probability of r successes in n trials for a given value of p is given by

$$f(D_{r,n}|p) = p^r (1-p)^{n-r} \tag{4}$$

This gives for the posterior distribution of p, after the data have been utilized,

$$f_F(p|D_{r,n}) = \frac{p^r (1-p)^{n-r}}{\int_0^1 p^r (1-p)^{n-r}\, dp} \tag{5}$$

This is the well known Beta Distribution. The most likely value is given by the maximum of $f_F(p|D_{r,n})$, which can be found as

$$\frac{df_F(p|D_{r,n})}{dp} = 0 \tag{6}$$

which gives the most likely value of p as

$$p = \frac{r}{n} \tag{7}$$

Then the 95 percent confidence limits, p_u and p_L, are given by

$$\int_{p_L}^{p_u} f_F(p|D_{r,n})dp = .95 \tag{8}$$

It can be shown that the Beta distribution for large n can be approximated by a normal distribution with a mean, $\mu(p)$ of r/n and a variance, $\sigma^2(p)$ of $\frac{r}{n}(1 - \frac{r}{n})/n$. Thus the normal tables can be used to find the 95 percent confidence limits p_u and p_L. For the normal approximation this is given by

$$p_L = \pm \frac{1.965}{\sqrt{n}} \sigma(p) \tag{9}$$

As shown by Cohen, in reference 20, the zero order, $P(F_o)$, first order, $P(F_1|F_o)$, and second order, $P(F_1|F_o\ F_{-1})$, probabilities can all be obtained from any four given empirical probability values. For example, Table 2-1 presents a set of calculated values based upon the empirical values for $P(F_o)$, $P(F_1|F_o)$, $P(F_1|F_o\ F_{-1})$, and $P(F_1|U_o\ U_{-1})$.

Some immediately derivable expressions are listed below as equations 10 through 14.

$$P(U_o) = 1 - P(F_o) \tag{10}$$

For a first order process,

$$P(U_1|F_o) = 1 - P(F_1|F_o) \tag{11}$$

and

$$P(F_1|U_o) = \frac{P(F_o)}{P(U_o)} P(U_1|F_o) \tag{12}$$

While for a second order process,

$$P(U_1|F_o\ F_{-1}) = 1 - P(F_1|F_o\ F_{-1}) \tag{13}$$

and

$$P(U_1|U_o\ U_{-1}) = 1 - P(F_1|U_o\ U_{-1}) \tag{14}$$

Similarly, it has been shown by Cohen (ref. 20) that

$$P(U_1|F_o\ U_{-1}) = \frac{P(F_o)\ [1-P(F_o|F_{-1})] - P(F_{-1})\ P(F_o|F_{-1})[1-P(F_1|F_o\ F_{-1})]}{P(F_o) - P(F_{-1})\ P(F_o|F_{-1})}$$

This can be rewritten as

$$P(U_1|F_o\ U_{-1}) = \frac{P(F_1)\ P(U_1|F_o) - P(F_1)\ P(F_1|F_o)\ P(U_1|F_o\ F_{-1})}{P(F_1)\ [1 - P(F_1|F_o)]}$$

$$= \frac{P(U_1|F_o) - P(F_1|F_o)\ P(U_1|F_o\ F_{-1})}{P(U_1|F_o)} \tag{15}$$

This is equivalent to

$$P(F_1|F_o\ F_{-1}) = \frac{P(F_1|F_o)\ P(U_1|F_o\ F_{-1})}{P(U_1|F_o)} \tag{16}$$

By interchanging F and U, one can also obtain

$$P(U_1|U_o\ F_{-1}) = \frac{P(U_1|U_o)\ P(F_1|U_o\ U_{-1})}{P(F_1|U_o)} \tag{17}$$

and equation (14) can be used to obtain

$$P(F_1|U_o\ U_{-1}) = 1 - P(U_1|U_o\ U_{-1}) \tag{18}$$

The remaining calculated values of Table 2-1 are complements of values found by equations (15) and (17). Differences between corresponding pairs of numbers in this table are but a few percent and they are always within the confidence limits. To illustrate the determination of confidence limits, suppose that $P(F_1|U_o)$ is to be computed from $P(F_1|F_o)$ by equations (11) and (12). Since $P(F_1|F_o)$ is a random variable, its 95-percent confidence limits p_L are given by equation (9) where $p = p(F_1|F_o)$. Using the symbols < > to denote the averaging process, the confidence limits of the calculated quantity, $P(U_1|F_o)$ are found by deriving the variance $\sigma^2[P(U_1|F_o)]$ from

$$\begin{aligned}\sigma^2[P(U_1|F_o)] &= \langle[1 - P(F_1|F_o)]^2\rangle - \langle[1 - P(F_1|F_o)]\rangle^2 \\ &= \langle P(F_1|F_o)^2\rangle - \langle P(F_1|F_o)\rangle^2 = \sigma^2[P(F_1|F_o)] \qquad (19)\end{aligned}$$

Since the variance is unchanged by subtraction, the confidence limits of the calculated value $P(U_1|F_o)$ is equal to the empirically-determined value found for $p(F_1|F_o)$.

Therefore, from Table 2-1 and equation (11),

$$\begin{aligned}P(U_1|F_o) &= (1 - 0.788) \pm 0.028 \\ &= 0.212 \pm 0.028 \qquad (20)\end{aligned}$$

From equation (12), $P(F_1|U_o) = K\, P(U_1|F_o)$ (21)

where

$$K = \frac{p(F)}{p(U)}$$

Following the same procedure as before,

$$\begin{aligned}\sigma^2[P(F_1|U_o)] &= \langle[K\, P(U_1|F_o)]^2\rangle - \langle[K\, P(U_1|F_o)]\rangle^2 \\ &= K^2\, \sigma^2[P(U_1|F_o)] \qquad (22)\end{aligned}$$

Thus the confidence limit is changed by a factor of K. Now from equations (20) and (21) and Table 2-1,

$$P(F_1|U_o) = \frac{0.694}{0.306}(0.212 \pm 0.028)$$
$$= 0.481 \pm 0.064 \qquad (23)$$

This result is shown in Table 2-1. Now since $p(F_1|U_c)$ and $p(F_1|F_o)$ can be found experimentally, either one can be calculated from the other. In making the choice of an empirical value, it is reasonable to select the value which, together with its corresponding calculated value, has the narrowest confidence limits. Such selections are marked with asterisks in Table 2-1 and are also plotted in Figure 2-2.

2.3.2 Analysis of the Order of the Markov Process

Figure 2-2 shows the probability of a favorable day for given conditions during the previous days. These results were taken from Table 2-1.

If the present case were described by a zero order Markov process, the conditional probabilities, $P(F_1|F_o)$ or $P(F_1|U_o)$, would be equal to $P(F_1)$. However, as shown in Figure 2-2 they are significantly different from the value of $P(F_1)$, so that it is extremely unlikely that a zero order Markov Process would describe the present case. It is also shown in Figure 2-2 that the probability of the next day being a favorable day, given that the present day is favorable, will also depend on the previous day; that is, $P(F_1|F_o\ U_{-1})$ is significantly different from $P(F_1|F_o\ F_{-1})$. This means that if the present day is favorable, the probability of the next day being favorable would depend on whether the previous day was favorable or unfavorable.

Thus the condition $P(\cdot|F\cdot)$ can be considered a second order process. However, in Figure 2-2 note that $P(F_1|U_o\ U_{-1})$ and $P(F_1|U_o\ F_{-1})$ are very close to $P(F|U)$. This means that $P(\cdot|U)$ can be considered a first order process. The results for the probability of an unfavorable day are equal to 1 minus the results for a favorable day and are also shown in Figure 2-2.

2.4 RUNS OF PERSISTENT ASTP LAUNCH CONDITIONS

In counting the number of days in a run, all sequences which originate before July 1 are excluded and those which extend into September are included, both arbitrarily.

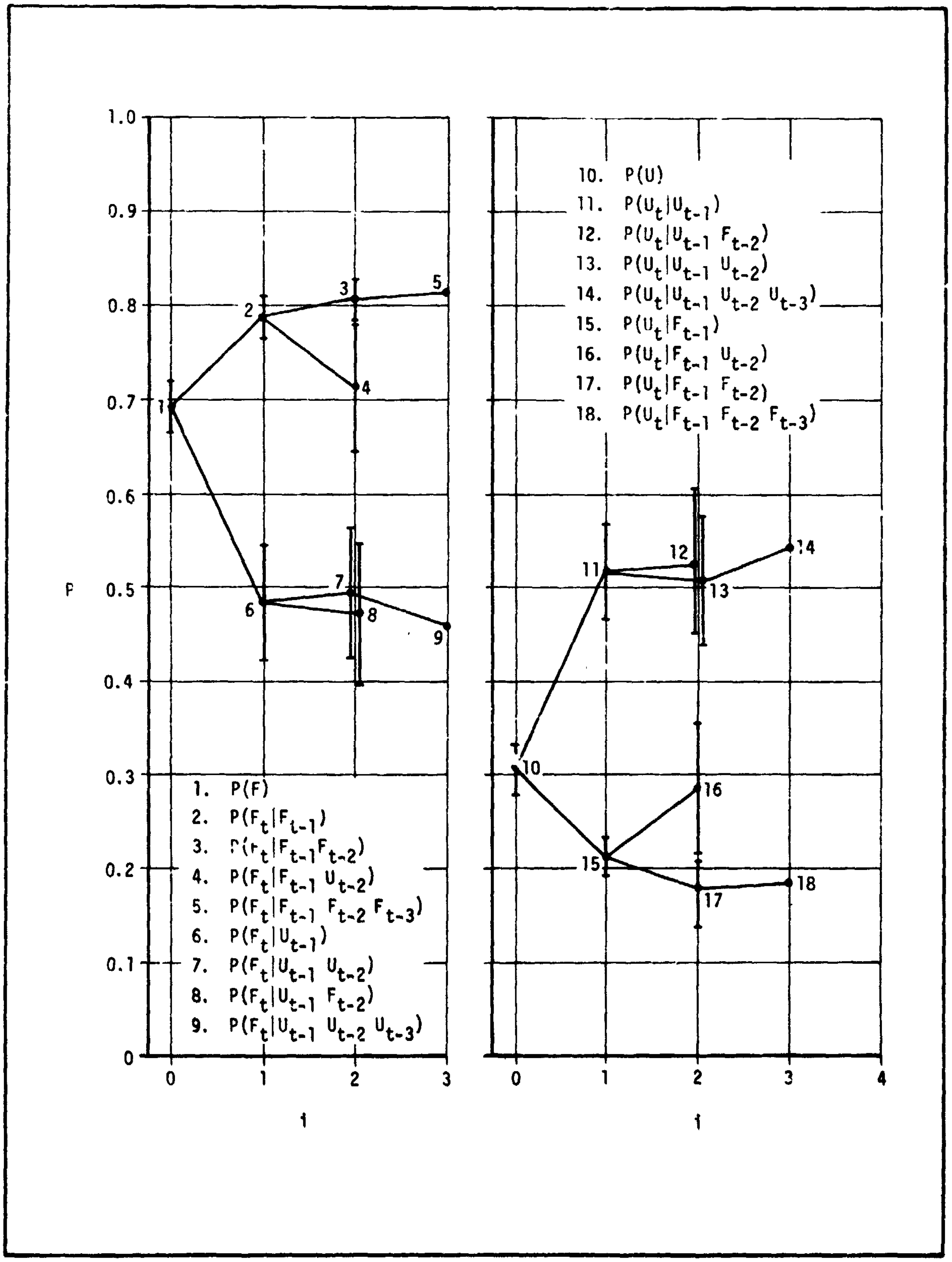


Figure 2-2. PROBABILITY (P) OF FAVORABLE (F) OR UNFAVORABLE (U) CONDITIONS THE NEXT DAY, GIVEN i DAYS OF RECORD. THE 95 PERCENT CONFIDENCE LIMITS ARE SHOWN FOR i $\leq$ 2

2.4.1 Runs of Unfavorable Days

The longest run of consecutive unfavorable days after a favorable day is one week long. This is shown in Figure 2-3, where relative frequencies give the empirical probability of runs of n or more days, with n ranging up to 7 days. Note that $p(n \geq 1)$ is the probability of one or more unfavorable days after a favorable day, and its empirical value of 0.212 is equal to $1-P(F_1|F_o)$ in Table 2-1.

The theoretical probabilities of the same runs are calculable from the basic empirical values found by counting cases. Thus the probability of n or more unfavorable days after a favorable day is given by

$$P(U^{(n)} F_{-1}) = P(F_{-1}) \; P(U_o|F_{-1}) \; P^{n-1}(U_1|U_o) \qquad (24)$$

The equations for a chain of events is discussed more fully in Section III. Normalizing yields

$$y_n = \frac{P(U^{(n)} F_{-1})}{P(F_{-1}) \; P(U_o|F_{-1})} = P^{n-1} \; (U_1|U_o) \qquad (25)$$

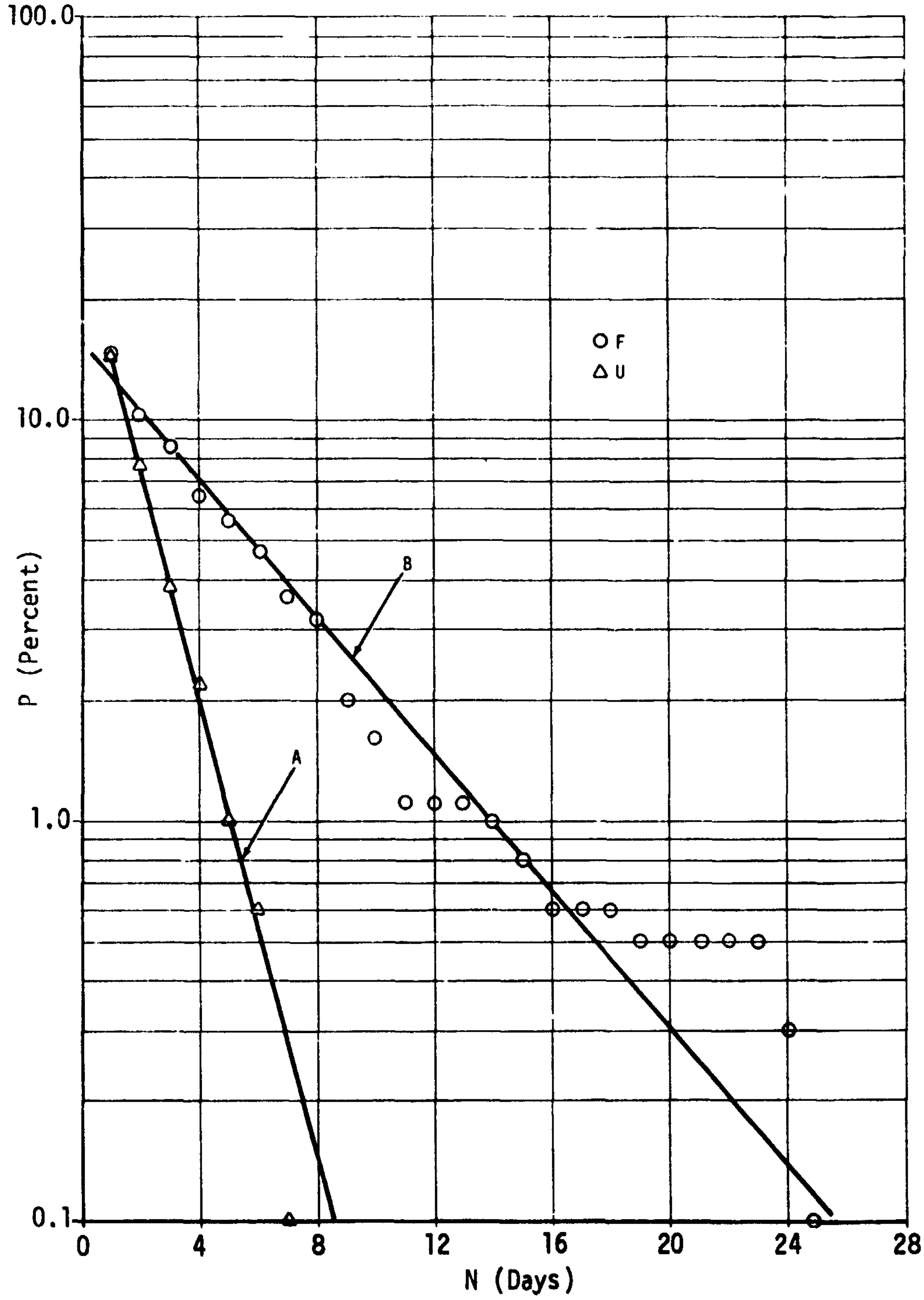
Curve A (Figure 2-3) shows the result when the probabilities are evaluated from the data by equation (24), and it is in fine agreement with the empirical information, at least to 0.1 percent. The utility of this graphical relationship is to read the probability of unfavorable ASTP launch weather for n or more days, given the present day is favorable.

Since Curve A is straight on semilog paper for $n \geq 1$, then

$$y_n = y_o \; e^{-k(n-n_o)} \qquad (26)$$

Evaluating k directly from Curve A yields $k = 0.658$. Then for a one-day advance,

$$\frac{y_{n+1}}{y_n} = P(U_1|U_o) = e^{-k(1)} = 0.518 \qquad (27)$$



EMPIRICAL VALUES OBTAINED BY COUNTING ARE ALSO SHOWN AS INDIVIDUAL POINTS

Figure 2-3. CUMULATIVE PROBABILITY (P) IN PERCENT OF THE OCCURRENCE OF N OR MORE UNFAVORABLE DAYS, AFTER A FAVORABLE DAY (CURVE A), EQ. (24), AND OF N OR MORE FAVORABLE DAYS, AFTER AN UNFAVORABLE DAY (CURVE B), BASED UPON 124 REVERSALS IN EACH CASE OUT OF A TOTAL OF 826 DAYS

This outcome for $P(U_1|U_o)$ agrees closely with the empirical value (0.517) obtained by counting cases. A similar linear relationship on runs was also examined by Langley in reference 21 for wet periods in Montreal.

2.4.2 Runs of Favorable Days

Using the same counting rule, the longest run of favorable conditions after an unfavorable day is found to be 25 days in length (Figure 2-3). Again, the probabilities have been estimated by relative frequencies counted from the data. The theoretical probabilities of n or more favorable days after an unfavorable day is given by

$$P(F^{(n)}\ U_{-1}) = P(U_{-1})\ P(F_o|U_{-1})\ P(F_1|F_o\ U_{-1})\ P^{n-2}(F_2|F_1\ F_o) \qquad (28)$$

Normalizing gives

$$y'_n = \frac{P(F^{(n)}\ U_{-1})}{P(U_{-1})\ P(F_o|U_{-1})} = P(F_1|F_o\ U_{-1})\ P^{n-2}(F_2|F_1\ F_o) \qquad (29)$$

Curve B (Figure 2-3) shows this result when the probabilities in equation (28) are evaluated from the data. Agreement with the plotted points is fairly good, although there are trends away from the theoretical curve when the probability falls below 0.02. Again, the probability of persistently favorable launch weather for n or more days, given the present day has unfavorable conditions, can be read from the graph.

Estimating k directly from Curve B yields a value of 0.203. Then for a one-day progression,

$$\frac{y'_{n+1}}{y'_n} = P(F_1|F_o\ F_{-1}) = e^{-k(1)} = 0.816 \qquad (30)$$

This outcome for $P(F_1|F_o\ F_{-1})$ is slightly less than the empirical value of 0.825 obtained by counting.

Section III

CONDITIONAL PROBABILITIES FROM THE CHAPMAN-KOLMOGOROV EQUATIONS

In this section, the applicability of the Markov chain theory will be stressed, and the Chapman-Kolmogorov Equations (Breiman, ref. 22) will be used to obtain conditional probabilities for up to 4 days following a base day with its identifiable favorable or unfavorable condition and available past weather record.

3.1 MARKOV CHAINS OF ORDER ONE AND ORDER TWO

The law of multiplication in probability theory states that the probability of a string of events E_t, E_{t+1}, ..., E_{t+n} is the product of n factors,

$$P(E_t, E_{t+1}, \dots E_{t+n}) = P(E_t)\, P(E_{t+1}|E_t)\, P(E_{t+2}|E_{t+1}, E_t) \dots P(E_{t+n}|E_{t+n-1}, \dots, E_{t+1}, E_t) \quad (31)$$

A Markov chain of first order also has n factors, but it retains but one previous event as "given" in each factor, thus,

$$P(E_t, E_{t+1}, \dots E_{t+n}) = P(E_t)\, P(E_{t+1}|E_t)\, P(E_{t+2}|E_{t+1}) \dots P(E_{t+n}|E_{t+n-1}) \quad (32)$$

Analogously, a Markov chain of second order retains two previous events as "given" in each factor, so that

$$P(E_t, E_{t+1}, \dots E_{t+n}) = P(E_t)\, P(E_{t+1}|E_t)\, P(E_{t+2}|E_{t+1}, E_t) \dots E(E_{t+n}|E_{t+n-1}, E_{t+n-2}) \quad (33)$$

Many of the papers referenced in Section I which treat persistence in precipitation and in dry periods find that a Markov model, especially the first order chain, is an acceptable device to describe the data.

3.2 THE CHAPMAN-KOLMOGOROV EQUATIONS

The probability of passing from the initial state A_i at time n to state A_j at time n+1 can be written as $P(A_{j,n+1}|A_{i,n})$. The probability of passing from the initial state A_i at time n to a new state A_k at time n+2 is given by the Chapman-Kolmogorov equation for a first order Markov process as

$$P(A_{k,n+2}|A_{i,n}) = \sum_j P(A_{k,n+2}|A_{j,n+1})\ P(A_{j,n+1}|A_{i,n}) \tag{34}$$

Thus the probabilities have been summed over all possible intermediate states A_j at time n+1. This process can be continued,

$$P(A_{\ell,n+3}|A_{i,n}) = \sum_k P(A_{\ell,n+3}|A_{k,n+2})\ \ P(A_{k,n+2}|A_{i,n}) \quad , \tag{35}$$

etc. The same type of analysis is applicable to second order or zero order processes.

This result can be used to predict unfavorable or favorable conditions as far ahead as desired. The results for favorable conditions can be obtained from the unfavorable relationships by the following equations.

$$P(F_i|U_o) = 1 - P(U_i|U_o) \tag{36}$$

$$P(F_i|F_o\ F_{-1}) = 1 - P(U_i|F_o\ F_{-1}) \tag{37}$$

$$P(F_i|F_o\ U_{-1}) = 1 - P(U_i|F_o\ U_{-1}) \tag{38}$$

These results can be used to predict the behavior in advance, given the present known weather conditions. The method of calculation is illustrated in Figure 3-1, and the outcome for U_i, i = 1, 2, 3, 4 for July and August at KSC is shown in Table 3-1 and Figures 3-1 and 3-2. Experimental counts are included in the table for comparison with the theoretical predictions, and agreement is to within 5 percent absolute value in all of the conditional probabilities. This indicates that the Chapman-Kolmogorov Equations are indeed applicable to this problem.

Table 3-1. CONDITIONAL PROBABILITIES OF UNFAVORABLE CONDITIONS FOR ASTP LAUNCH

$P(U_i\|U_o)$			$P(U_i\|F_o\ U_{-1})$		$P(U_i\|F_o\ F_{-1})$	
ith day	E*	C**	E*	C**	E*	C**
1	0.517	0.517	0.272	0.216	0.175	0.193
2	0.410	0.454	0.267	0.249	0.242	0.255
3	0.331	0.372	0.319	0.270	0.274	0.284
4	0.306	0.343	0.245	0.278	0.278	0.297

**Empirical value obtained by counting and use of the complementary relationship.*

***Calculated value from the Chapman-Kolmogorov Equations.*

$P(U_i | U_o)$ - probability of the i^{th} unfavorable day given that today is unfavorable.

$P(U_i | F_oU_{-1})$ - probability of the i^{th} unfavorable day given that today is favorable and yesterday unfavorable.

$P(U_i | F_oF_{-1})$ - probability of the i^{th} unfavorable day given that today is favorable and yesterday favorable.

Unfavorable conditions mean that one or more of the four launch constraints (see page 2-1) occurred at least once between 1200 EST and 1700 EST.

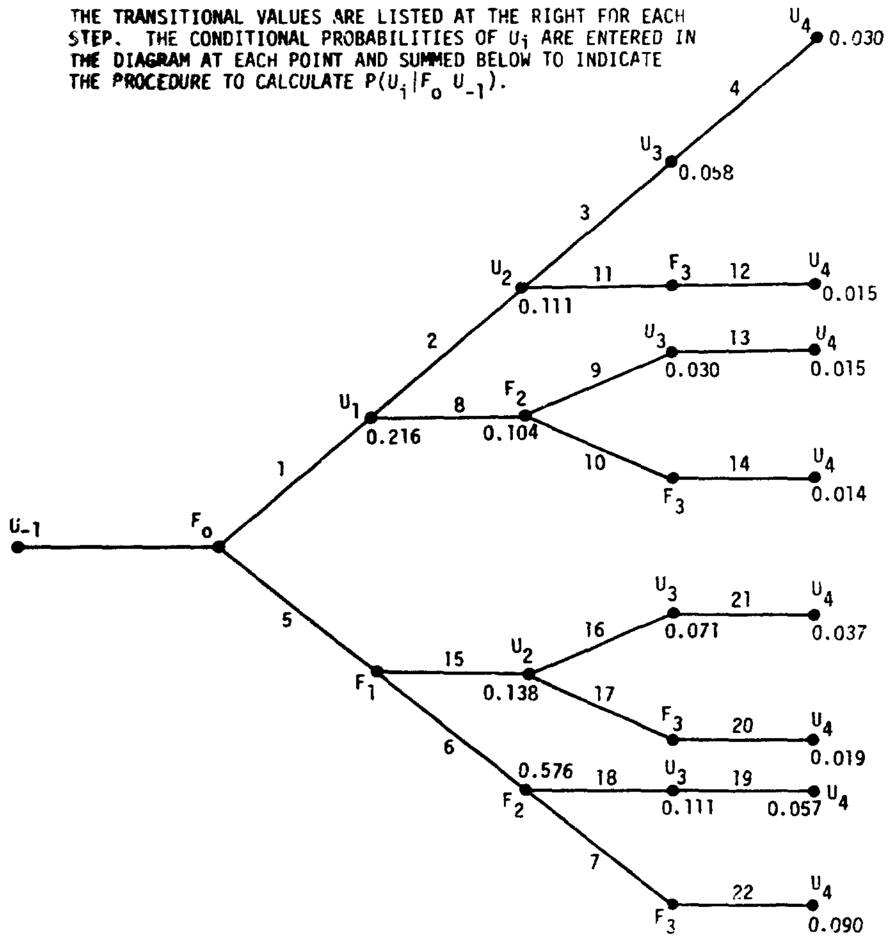
Examples illustrating use of the tables:

1. This afternoon was unfavorable for ASTP launch. The empirical and calculated probability that tomorrow afternoon will be unfavorable is .517. The empirical probability that the third afternoon from today will be unfavorable is 0.331.

2. This afternoon was favorable, but the previous afternoon was unfavorable. The calculated probability that tomorrow afternoon will be unfavorable is 0.216. The empirical probability that the 4th afternoon from today will be unfavorable is 0.245.

3. This afternoon was favorable and the previous afternoon was favorable also. The empirical probability that tomorrow afternoon will be unfavorable is 0.175. The calculated probability that the afternoon after tomorrow (2nd day in the future) will be unfavorable is 0.255.

THE TRANSITIONAL VALUES ARE LISTED AT THE RIGHT FOR EACH STEP. THE CONDITIONAL PROBABILITIES OF U_i ARE ENTERED IN THE DIAGRAM AT EACH POINT AND SUMMED BELOW TO INDICATE THE PROCEDURE TO CALCULATE $P(U_i|F_o\ U_{-1})$.



i	1	2	3	4
$P(U_i\|F_o\ U_{-1})$	0.216	0.249	0.270	0.278

1. $P(U_1|F_o\ U_{-1}) = 0.216$
2. $P(U_2|U_1) = 0.517$
3. $P(U_3|U_2) = 0.517$
4. $P(U_4|U_3) = 0.517$
5. $P(F_1|F_o\ U_{-1}) = 0.714$
6. $P(F_2|F_1\ F_o) = 0.807$
7. $P(F_3|F_2\ F_1) = 0.807$
8. $P(F_2|U_1) = 0.483$
9. $P(U_3|F_2\ U_1) = 0.286$
10. $P(F_3|F_2\ U_1) = 0.714$
11. $P(F_3|U_2) = 0.483$
12. $P(U_4|F_3\ U_2) = 0.286$
13. $P(U_4|U_3) = 0.517$
14. $P(U_4|F_3\ F_2) = 0.193$
15. $P(U_2|F_1\ F_o) = 0.193$
16. $P(U_3|U_2) = 0.517$
17. $P(F_3|U_2) = 0.483$
18. $P(U_3|F_2\ F_1) = 0.193$
19. $P(U_4|U_3) = 0.517$
20. $P(U_4|F_3\ U_2) = 0.286$
21. $P(U_4|U_3) = 0.517$
22. $P(U_4|F_3\ F_2) = 0.193$

Figure 3-1. BRANCHING DIAGRAM FOR THE COMPUTATION OF CONDITIONAL PROBABILITIES BY THE CHAPMAN-KOLMOGOROV EQUATIONS

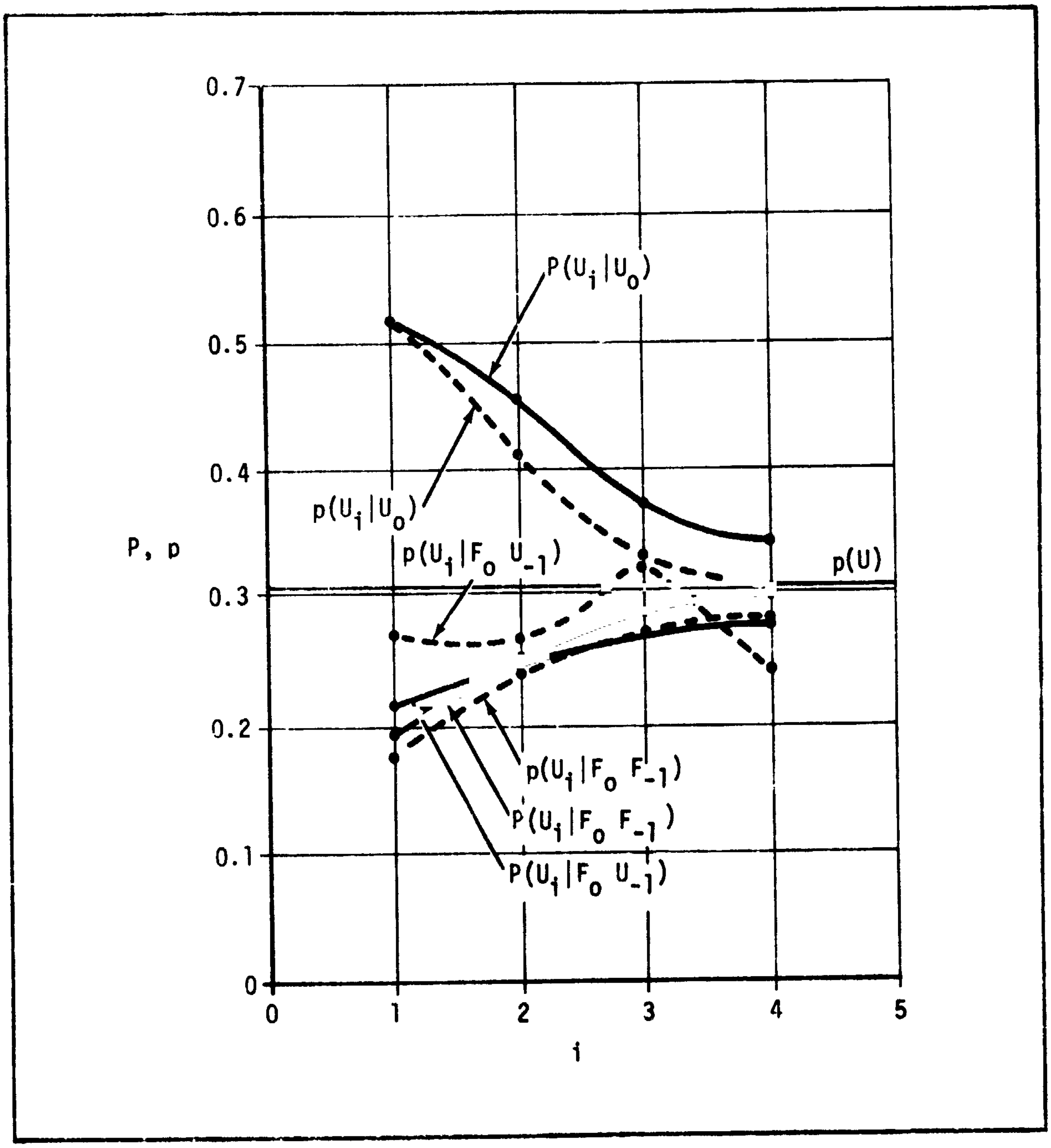


Figure 3-2. CONDITIONAL PROBABILITIES OF MARGINAL ASTP LAUNCH CONDITIONS IN JULY AND AUGUST AT CAPE KENNEDY, BASED UPON THE CHAPMAN-KOLMOGOROV EQUATIONS (P) AND UPON EMPIRICAL COUNTS (p), FOR i DAYS IN THE FUTURE FROM DAY ZERO.

Section IV

EXAMPLES OF PROBABILITY CALCULATIONS

The values of Table 2-1 and their complements for unfavorable cases are sufficient to compute the probabilities of runs with no previous conditions. For example, if $P(U_oU_1U_2)$ is desired, equation (32) is applied to get

$$P(U_oU_1U_2) = P(U_o)\ P(U_1|U_o)^2 = (0.306)\ (0.517)^2 = 0.082$$

If $P(F_oF_1F_2)$ is desired, equation (33) is applied to yield

$$P(F_oF_1F_2) = P(F_o)\ P(F_2|F_1F_o)^2 = (0.694)(0.823)^2 = 0.470$$

In case a probability specified at a particular hour is sought, the appropriate link from subsection 2.1 is inserted. Thus, to find the 1500 EST probability of a restriction, given a previous day's restriction at 1500 hours, use

$$\begin{aligned} P(U_{1500,1}|U_{1500,0}) &= P(U_o|U_{1500,0})\ P(U_{1500,1}|U_o) \\ &= P(U_o|U_{1500,0})\ P(U_1|U_o)\ P(U_{1500,1}|U_1) \\ &= (1.00)\ (0.517)\ (0.434) = 0.224 \end{aligned}$$

The unconditional probability of encountering two consecutive restrictions at 1500 hours is

$$\begin{aligned} P(U_{1500,1}\ U_{1500,0}) &= P(U_{1500,0}|U_o)\ P(U_o)\ P(U_1|U_o)\ P(U_{1500,1}|U_1) \\ &= (0.306)\ (0.434)\ (0.517)\ (0.434) = 0.030 \end{aligned}$$

Section V

SUMMARY AND CONCLUSIONS

In the Apollo-Soyuz Test Program, the decision to launch must be made at least 15 hours before launch time. Therefore, mission planning can be aided by a knowledge of statistical relationships between the occurrence of inclement (and favorable) ASTP weather conditions in past time and future time. Markov theory has been applied in this study to elucidate these relationships in terms of conditional probabilities for Kennedy Space Center.

The first forecasting problem investigated was the length of record of past weather which is useful to a prediction. Based upon the historical sequence of hourly reports for July and August from 1957 to 1971, relative frequencies of marginal ASTP weather were gleaned from the data and expressed as four empirical conditional probabilities from which other conditional probabilities up to second order were derived. The outcome is contingent upon the nature of the preceding weather. Thus, if afternoon weather for the current day has been unfavorable, the previous afternoon's reports have negligible forecast value. On the other hand, if the afternoon weather for the current day has been favorable, the previous day's reports are important to the prediction. These results signify that first order and second order Markov chains, respectively, are operative.

The second forecasting problem studied was the matter of runs of favorable or unfavorable launch conditions. Such runs were found to persist as long as 25 days and 7 days, respectively. In the case of unfavorable present weather, there is a conditional probability slightly greater than 0.50 that inclement conditions will persist another day. On the other hand, the probability of one or more favorable days, given unfavorable present weather, is only about 0.15. A probability can be read from Figure 2-3 for any desired number of days in a sequence of favorable days after an unfavorable day.

Figure 2-3 is also useful to an analysis of runs of unfavorable weather. If one assumes the present weather is good, for example, then the probability of a change to inclement weather for one or more days is 0.14, approximately. Conditional probabilities of runs of unfavorable days with greater minimum length are readily obtained from this graph.

The final forecasting problem investigated was the prediction of ASTP launch conditions for a few days ahead, following a base day with known present weather and past weather. Further application of Markov theory in the form of the Chapman-Kolmogorov Equations was made, there being evidence of feasible predictions up to four days, at which time the empirical value of $p(U_4|U_0)$ generally reaches the unconditional value of $p(U)$. These results are available in tabular form (Table 3-1) and also graphical form (Figures 3-1, 3-2).

The theoretical results have been compared with the experimental counts to determine the amount of agreement between Markov chains of the order indicated by preliminary investigation of the first forecasting problem, and the data. Agreement was within 5 percent absolute value in all of the computations (Table 3-1).

Finally, it is noted again that the definition of "unfavorable" used throughout this study is "the occurrence of one or more of the current ASTP constraints (subsection 2.1) at any hour between 1200 EST and 1700 EST". There is a known probability that an ASTP restriction will take place at a particular hour on an unfavorable day in July and August, and as was demonstrated in subsection 2.5, this provides a link between the tabular values of this report and calculations of probabilities for specific hours.

Section VI

REFERENCES

1. Feyerherm, A. M. and Bark, L. Dean, "Statistical Methods for Persistent Precipitation Patterns", Journ. Appl. Meteor., 4, 1965, 320-328.

2. Williams, C. B., "Sequences of Wet and of Dry Days Considered in Relation to the Logarithmic Series", Quart. J. R. Meteor. Soc., 1952, 91-96.

3. Hopkins, J. W. and Robillard, P., "Some Statistics of Daily Rainfall Occurrence for the Canadian Prairie Provinces", Journ. Appl. Meteor., 3, 1964, 600-602.

4. Weiss, Leonard L., "Sequences of Wet or Dry Days Described by a Markov Chain Probability Model", Mon. Wea. Rev., 92, 1964, 169-176.

5. Eriksson, Bertil, "A Climatological Study of Persistency and Probability of Precipitation in Sweden", Tellus, 17, 1965, 484-497.

6. Brelsford, William M. and Jones, Richard H., "Estimating Probabilities", Mon. Wea. Rev., 95, 1967, 570-576.

7. Smith, C. G., "Monthly, Seasonal and Annual Fluctuations of Rainfall at Oxford since 1815", Weather, 29, 1974, 2-16.

8. Gabriel, K. R. and Neumann, J., "On a Distribution of Weather Cycles by Length", Quart. J. R. Meteor. Soc., 83, 1957, 375-380.

9. Gabriel, K. R. and Neumann, J., "A Markov Chain Model for Daily Rainfall Occurrence at Tel Aviv", Quart. J. R. Meteor. Soc., 88, 1962, 90-95.

10. Caskey, Jr., James E., "A Markov Chain Model for the Probability of Precipitation Occurrence in Intervals of Various Length", Mon. Wea. Rev., 91, 1963, 298-301.

11. Gabriel, K. R., "The Distribution of the Number of Successes in a Sequence of Dependent Trials", Biometrika, 46, 1959, 454-460.

12. Walker, Strother H. and Duncan, David B., "Estimation of the Probability of an Event as a Function of several Independent Variables", Biometrika, 54, 1967, 315-327.

13. Von Mises, R., "Mathematical Theory of Probability and Statistics", New York, Academic Press, 1964, 184-202.

14. Feller, W., "An Introduction to Probability Theory and its Applications", 2nd Ed., New York, John Wiley and Sons, Inc., 1957, Ch. XII.

15. Wilks, S. S., "Mathematical Statistics", New York, John Wiley and Sons, Inc., 143-144.

16. Anderson, T. W. and Goodman, L. A., "Statistical Interfeience about Markov Chains", Ann. Math. Stat., 28, 1957, 89-110.

17. Falls, Lee W., "A Probability Distribution for the Number of Thunderstorm Events at Cape Kennedy, Florida", NASA TMX-53816, February 20, 1969, 19 pp.

18. Neumann, Charles J., "The Thunderstorm Forecasting System at Kennedy Space Center", Journ. Appl. Meteor., 10, 1971, 921-936.

19. Pratt, John W., Raiffa, Howard, and Schlaifer, Robert, "Introduction to Statistical Decision Theory", New York, McGraw-Hill Book Company, 1965, Ch. 11.

20. Cohen, A. [illegible]ford, "The Use of Second Order Markov Chains in Estimating Probabilities of Sequences of Favorable and Unfavorable Days for Launching Space Vehicles", Technical Report Number 29, Institute of Statistics, The U. of Georgia, July 25, 1968, 13 pp.

21. Langley, R. W., "The Length of Dry and Wet Periods", Quart. J. R. Meteor. Soc., 79, 1953, 520-527.

22. Breiman, Leo, "Probability and Stochastic Processes", Boston, Houghton Mifflin Company, 1969, Ch. 6, 7.

NASA—MSFC