دانشگاه پیام نورمرکز تهران

دانشکده علوم ریاضی

پایان نامه کارشناسی ارشد

پیش بینی بر اساس مدل های احتمال

استاد راهنما :

دکتر پرویز نصیری

استاد مشاور :

دکتر علی شادرخ

گردآورنده :

توماج یوسف زاده

تابستان ۸۷

بسم الله الرحمن الرحيم

تقدیم به:

مادرم

که نمونه عشق، محبت و فداکاری

پدرم

که نمونه پشتکاری، تلاش و استواری

خانواده ام

که مشوق من در زندگی

و

همسر عزیزم

که نمونه دوست داشتن، صمیمیت، سادگی، مهربانی و دلسوزی است.

تقدیر و تشکر

سپاس بی کران خداوند حکیم را ، که به من حقیر توان موفقیتی دوباره عطا فرمود.

ضمن عرض خسته نباشید حضور تمامی اساتید و دست اندرکاران که با حضور و تلاش روزافزون خود در پیشرفت و نشر علم و دانش نقشی سازنده دارند، اینجانب وظیفه خود می دانم که از زحمات بی دریغ استادان گرانقدرم خصوصا جناب آقای دکتر پرویز نصیری، دکتر علی شادرخ و اساتید داور آقایان دکتر مسعود گنجی و دکتر باقر مقدس زاده نهایت تشکر و قدر دانی را داشته باشم.

توماچ یوسف زاده

فهرست مطالب

عنوان **صفحه**

فهرست جداول

چکیده

آنالیز نوبت بارندگی، قسمت مهمی از آنالیز داده های بارندگی است. طـول نوبـت بارنـدگی، اطلاعات مفیدی در خصوص زمان توزیع بارندگی را ارائه می دهد. آنـالیز آمـاری طـول نوبـت هـای بارندگی بر آن است که بهترین مدل آماری را برای الگو های طـول نوبـت مـشاهده ارائـه دهـد، بـه طوریکه مدل ارائه شده ممکن است برای مکان های مختلف، متفاوت باشـد. طـول نوبـت بارنـدگی، دوره ای از روز های بارانی است که از یک روز خشک شروع می شود و به یک روز خشک ختم می شود. در این پایان نامه چندین مدل آماری برای طول نوبت بارندگی بحث می شود: (۱)مـدل گلـد، (۲)مدل ککران، (۳)مدل مارکوف ککـران، (۴)مـدل سـری هـای لگـاریتمی، (۵)مـدل سـری هـای لگاریتمی مارکوف. مدل گلد، مدل ککران و مدل سری های لگاریتمی بر اساس استقلال شرایط آب و هوایی از یک روز به روز دیگر به دست می آیند، امـا مـدل هـای مـارکوف ککـران و سـری هـای لگاریتمی مارکوف تحت فرض وابستگی مارکوف حاصل می شوند.

فصل ۱

مقدمه

یکی از بلایای مهم تولیدات کشاورزی پدیده تگرگ است . میزان خـسارت تگـرگ بـسته بـه نـوع محصول، اندازه تگرگ ، شدت رگبار آن و مرحله فنولوژیکی گیاه متفـاوت اسـت. دسـتیابی بـه روش هـای مقابله با این پدیده خسارت زا ، در گرو مطالعه و شناخت وقـوع ایـن پدیـده و عوامـل بوجـود آورنـده آن و شرایط تکوین و تاثیر آن در مقیاسهای زمانی و مکانی معین است[۶].

تحمیل خسارت ناشی از تگرگ بر اقشار مختلف همه ساله در نقاط مختلف کشور اتفـاق مـی افتـد . هـدف کلی این پایانامه، بررسی و انتخاب بهترین مدل آماری در مورد پدیده تگرگ می باشد تـا بـا بهـره‌گیـری از نتایج این مطالعات ، روشهای فنی مقابله بـا تگـرگ و شـیوه هـای کـاهش خـسارت ناشـی از آن از طـرف دستگاههای ذیربط مورد برنامه ریزی قرار گیرد. همچنین با شناسـایی منـاطق آسـیب پـذیر کـشاورزی در مقابل این پدیده و بهره گیری از فن آوریها و تدابیر هواشناسی کشاورزی می توان به امـر کـاهش خـسارت تگرگ مبادرت ورزید[۳].

به منظور کاهش تگرگ در مناطق آسیب پذیر در قالب طرح تحقیقاتی مطالعه ای صـورت گرفتـه است [۶]. در این طرح با تعیین تعداد روزهای همرا با تگـرگ در ماههـا وفـصول مختلـف سـال بـه کمـک

روشهای گوناگون آماری سعی در انتخاب مدل آماری مناسب شده ، تا بهترین مدل گزینش شود ، البته بررسی های جدید تری در این زمینه انجام شده که مدل مناسب تر از میان دو مدل ککران[1] و مارکوف ککران[2]، مدل ککران است.

فصل ۲ مباحثی در موردمفاهیم کلی مدل بندی، تاریخچه و نمونه هایی از مدل های آماری می باشد که در مدل بندی روزهای بارانی و خشک مورد استفاده قرار گرفته تا یک تصویر و پیش زمینه ای از مدل بندی مورد نظر در این پایان نامه ایجاد شود.

فصل ۳ شامل بررسی مدل های گلد[3]، ککران و مارکوف ککران و معرفی توزیع سری های لگاریتمی[4] و توزیع سری های لگاریتمی مارکوف[5] می باشد .

فصل ۴ به برازش دو مدل ککران و مارکوف ککران به روی داده های شهر تبریز و مقایسه آن با مدل های برازش شده قبلی پرداخته [۶] و نتایج و پیشنهادات ارائه گردیده است.

[1]- Cochran model
[2]- Markov Cochran model
[3]-Gold model
[4]- Logarithmic series distribution
[5]- Markov logarithmic series distribution

فصل ۲

مدل بندی و مروری برتاریخچه مدل بندی روزهای خشک و بارانی

۲- ۱ مدل بندی

آمار شاخه ای از علم است که لزوما تمامی نظریه های آن در قالب ریاضی قابل بیان نیست و برخلاف اکثر شاخه های ریاضی که ابتدا موضوعی اثبات می شود و سپس در صورت امکان به دنبال کاربرد آن می روند ، در آمار غالبا پدیده ای مشاهده می شود و سپس برای توضیح و کنترل آن ، نظریه مربوطه بنا می گردد. گسترش و شاخ و برگ دادن به نظریه های آماری ، کاربرد آن را مستدل تر کرده ، از به انحراف کشیدن نتایج تحقیقات جلوگیری می نماید[۹۳] و[۷]. یکی از اساسی ترین روش های دستیابی به مطالب جدید علمی ، افزایش سطح توانائی در مدل سازی پدیده های مورد مطالعه است .

در مدل سازی ، یک فرآیند طبیعی تبدیل به الگوئی می گردد که قابل تجزیه وتحلیل از طریق تکنیکهایی است که می فهمیم و به آنها اعتماد داریم. یک مدل به مدلساز کمک می کند تا رفتار یک پدیده یا واقعه را پیش بینی کند یا توضیح دهد. مدلساز تعاریف و فرض هایی ساده کننده ارائه می دهد و

سپس تعدادی قوانین و اصول طبیعی که رفتار پدیده را به طور ملموس کنترل می کنند یا توضیح می دهند ، انتخاب می کند . مدل ساز با استفاده از قوانین طبیعی ، متغییر ها ، مفاهیم ریاضی و ابزارهایی چون نرم افزار یک مدل ریاضی را می سازد . هدف مدل ساز این است که مدلی ارائه کند که به حد کافی عمومی باشد که رفتار پدیده را به طور ملموس توضیح دهد ولی آنقدر پیچیده نباشد که تجزیه تحلیل مدل را غیر ممکن سازد . برای اطمینان به مدل ، مدل ساز موارد خاص را با مدل حل می کند و سپس نتایج به دست آمده را با نتایج حاصل از آزمایشات مقایسه می کند ، از این طریق مدل ساز به حدود کاربرد مدل پی می برد و این محدودیت را با عنوان حدود اعتبار[1] مطرح می سازد[۹۳].

۲-۲ مروری بر تاریخچه مدل بندی روزهای خشک و بارانی

بطور کلی تجزیه و تحلیل طول نوبت بارندگی ، اطلاعات مفیدی در خصوص زمان توزیع بارندگی ارائه می دهد و از دیر باز مورد توجه پژوهشگران نظیر گلد[2] (۱۹۲۹) ، ککران[3] (۱۹۳۸) ، فیرم[4] و بارک[5] (۱۹۶۵) ، مدهی[6] (۱۹۷۶) ، چوداری[7] (۱۹۸۱) ، گوره[8] ونصیری[9] (۱۹۹۷) بوده است. [۱۳] و [۳۳] و [۳۴] و [۳۹] و [۴۰] و [۴۱] و [۴۲] و[۴۳] و[۴۵] و[۶۸].مسائل مربوط به برآورد توزیع احتمال طول دنباله روزهای خشک و بارانی به یک سری از تحقیقات نیم قرن اخیر بر می گردد. ویلیامز[10] (۱۹۵۲) در این

1- Regions of validity
2-Gold
3-Cochran
4- Feyerhim
5- Bark
6- Medhi
7- Chowdhury
8- Gore
9- Nasiri
10- Williams

زمینه پژوهش هایی انجام داد . وی توزیع هندسی را برای مدل بندی روزهای خشک و بارانی به کار برد و به توزیع سری لگاریتمی جهت تشریح و تفسیر فراوانی های وقوع روزهای خشک (و بارانی) با طول های متفاوت اشاره کرد[۱۳]. همچنین لانگلی[۱](۱۹۵۳) دنباله روزهای خشک و بارانی را در چند ایستگاه هواشناسی مورد مطالعه قرار داد و به این نتیجه رسید که احتمال یک روز خشک که روز قبل نیز خشک بوده ، با طول دوره خشک کمی افزایش می یابد [۱۳] و [۶۷]. (این مطلب به وابستگی روزهای خشک و بارانی اشاره دارد.)

گبریل[۲] و نیومن[۳] (۱۹۵۷) این مسئله را بسط داده و توزیع احتمال را برای چرخه هوایی (یعنی دنباله ای از روزهای خشک که پشت سر دنباله ای از روزهای بارانی واقع شده اند) با فرض مستقل بودن اعضای دنباله (روزهای خشک و بارانی) بدست آورده که در روش آنها ، طول نوبت بارندگی از توزیع هندسی پیروی می کرد[۱۳].

در یک مقاله دیگر گبریل (۱۹۵۹) نشان داد که ، توزیع احتمال دقیق تعداد موفقیت ها در یک دنباله از آزمایشهای وابسته ، می تواند در نتیجه گیری و انتخاب نهایی توزیع احتمال تعداد روزهای بارانی در یک دوره (دنباله) n روزه مورد استفاده قرار گیرد[۱۳].

به طور کلی روش هایی که جهت مدل بندی شرایط آب و هوایی نظیر روزهای بارانی و خشک ، روزهای طوفانی و..... به کار رفته ، عموما مدل های مربوط به فرآیند های تصادفی (به ویژه مدل زنجیره

---

[1]- Longley
[2]- Gabriel
[3]- Neumann

مارکف مرتبه اول) و مدل های سری زمانی است ، نظیر: ظفـر علـی[۱] (۱۹۷۱) ، فرمـور[۲](۱۹۷۱) ، کـوک[۳] (۱۹۵۳) ، ویس[۴](۱۹۶۴) ، چتفیلد[۵] (۱۹۶۴).

---

[1]- **Zaffar Ali**
[2]- **Fermor**
[3]-**Cooke**
[4]- **Weiss**
[5]- **Chatfield**

فصل ۳

بررسی مدل بندی روزهای بارانی و خشک

۳- ۱ مفاهیم اولیه

اگر X_1، X_2، ...، X_m دنباله ای از روزهـا در فـصل بارنـدگی باشـد ، بـرای هـر روز خـاص X_i احتمال ممکن وجود دارد که آن روز بارانی باشد یا نباشد . برای مثال فرض کنید بخواهیم بارانی بـودن یـا نبودن هر روز را در فصل بارندگی مورد مطالعه قرار دهیم.

بارانی بودن یا نبودن روز خاص بستگی به میـزان بارنـدگی دارد. معمـولا اداره هواشناسـی میـزان نزولات آسمانی را از ساعت ۸/۳۰ صبح تا ۸/۳۰ صبح روز بعد اندازه گیری مـی کنـد . اگـر میـزان نـزولات آسمانی بیشتر یا مساوی ۰/۱ میلی متر باشد آن روز را مرطـوب (بـارانی) ، وگرنـه آن روز را خـشک (غیـر بارانی) محسوب می کند[۵].

لازم به توضیح است که در بعضی از کشورها اگر میزان بارندگی در یک روز بیشتر یا مساوی ۲/۵ میلی متر باشد ، آن روز را مرطوب حساب می کنند[۳۳].

در این پایان نامه ، تعداد روزهای فصل بارندگی، m روز است که به صورت $X_1 ، X_2 ، ... ، X_m$ یا $\{X_i ، i \geq 1\}$ نمایش داده شده است که در آن $X_i = 1$ نشانگر آن است کـه آن روز بـارانی یـا میـزان بارندگی بیشتر یا مساوی ۲/۵ میلی متر ، $X_i = 0$ نشانگر عدم بارانی بودن یا میزان بارندگی کمتر از ۲/۵ میلی متر بوده است.

بر اساس این طبقه بندی ، طول نوبت بارندگی ، دوره ای از روزهای بـارانی اسـت کـه از یـک روز خشک شروع شده و به یک روز خشک ختم می شود . بر اساس طول نوبت بارندگی دو مدل آماری ، مـدل ککران و مدل مارکوف ککران معرفی و تشریح می شـوند کـه مـدل اول بـر اسـاس اینکـه روزهـا در فـصل بارندگی از هم مستقل اند ارائه می شود یعنی بارانی بودن یا نبودن روز قبل تاثیری در بـارانی بـودن امـروز ندارد.

ولی مدل مارکوف ککران ، با فرض وابستگی مرتبه اول بین روزها در فـصل بارنـدگی اسـت ، یعنـی بـارانی بودن یا نبودن امروز متاثر از بارانی بودن یا نبودن روز قبل است ، ارائه می شود.

۳-۲ مدل گلد

گلد در سال ۱۹۲۹ برای دنباله ای از روزهای $X_1 ، X_2 ، ... ، X_m$ در فـصل بارنـدگی ، بـا فـرض اینکه X_i ها از هم مستقل اند یا به عبارت دیگر ، بارانی بودن یا نبودن امروز مـستقل از روز قبـل اسـت و احتمال اینکه روز خاص بارانی باشد یا نباشد و دارای شانس یکـسان و برابـر بـا $\frac{1}{2}$باشـد ، مـدل زیـر بـرای محاسبه طول نوبت بارندگی مورد انتظار به طول r ارائه می داد :

$$f(r) = \frac{m - r + 3}{2^{r+2}} \qquad r = 1, 2, \ldots, m - 1 \qquad (۳-۲-۱)$$

۳-۳ مدل ککران

ککران در سال ۱۹۳۸ با بسط مدل گلد و با فرض اینکه شانس بارانی بودن یا نبودن در روز به خصوص مساوی نیست و دارای احتمالات $P[X_i = ۱] = p$ و $P[X_i = ۰] = q$ هستند (بطوریکه $p + q = ۱$) مدلی برای محاسبه طول نوبت بارندگی مورد انتظار به طول r ارائه داد که مراحل آن به صورت زیر می باشد:

الف) اگر شروع بارندگی ، روز اول فصل بارندگی باشد و r روز متوالی ادامه یابد یا روزهای بارندگی r روز آخر فصل بارندگی باشد، احتمال داشتن r روز متوالی برابر است با :

$$p^r q + q p^r = ۲ p^r q \qquad (۳-۳-۱)$$

ب) اگر شروع بارندگی غیر از روز اول یا ختم آن غیر از روز آخر باشد . به عبارت دیگر شروع آن روز ۲ یا ۳ یا ... $(m-r)$ باشد. تعداد حالات متمایز برابر با $m - r - ۱$ خواهد بود که احتمال هر حالت برابر است با: $p^r q^۲$ ، لذا طول نوبت بارندگی مورد انتظاردر این حالت برابر است با :

$$(m - r - ۱) p^r q^۲ \qquad (۳-۳-۲)$$

با توجه به روابط (۳-۳-۱) و (۳-۳-۲) طول نوبت بارندگی مورد انتظار به طول r برابر است با :

$$f(r) = \begin{cases} ۲ p^r q + (m - r - 1) p^r q^۲ & , r = ۱,۲,...,m-r \\ p^r & , r = m \end{cases} \qquad (۳-۳-۳)$$

مدل ککران با فرض $P[X_i = ۰] = P[X_i = ۱] = \frac{۱}{۲}$ همان مدل گلد است.

اگر r طول نوبت بارندگی در یک دورهٔ m روزه باشد ، می توان فراوانی مورد انتظار طول نوبت بارندگی با طولهای ۱ ، ۲ ، ۳ ، ... را از رابطهٔ (۳-۳-۳) بدست آورد.

اگر O_r ها فراوانی مشاهدات طول نوبت بارندگی با طول r ، و E_{rm} ها فراوانی مورد انتظار متناظر باشند ، انگاه جدول فراوانی مورد انتظار را می توان به صورت جدول (۳. ۳. ۳) نمایش داد.

......	۳	۲	۱	r
......	O_3	O_2	O_1	O_r
......	E_{3m}	E_{2m}	E_{1m}	E_{rm}

جدول۳. ۳. ۳ : فراوانی های مشاهده شده و مورد انتظار

۳-۴ مدل مارکوف ککران

گبریل و نیومن در سال ۱۹۶۲ و مدهی در سال ۱۹۶۷ نشان دادند که روزهای بارانی وغیر بارانی در طول فصل بارندگی مستقل نیستند [۴۱] و [۴۲]. بلکه بارانی بودن یا نبودن روز خاص، بستگی دارد به روز قبل که بارانی باشد یا نباشد. با فرض وجود وابستگی بین روزهای فصل بارندگی دنباله $\{X_m, m \geq 1\}$ دارای فضای نمونه $S=\{0,1\}$ و ماتریس تغییر وضعیت P است که عبارت است از :

$$P=\begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix}$$

به طوریکه :

$$\alpha = P[X_{n+1} = ۱ \mid X_n = ۰]$$

$$\beta = P[X_{n+1} = ۰ \mid X_n = ۱]$$

در واقع پارامترها به صورت زیر تعریف می شوند:

α: احتمال اینکه امروز بارانی باشد به شرط آنکه روز قبل بارانی نباشد.

β: احتمال اینکه امروز بارانی نباشد به شرط آنکه روز قبل بارانی باشد.

$۱-\alpha$: احتمال اینکه امروز بارانی نباشد به شرط آنکه روز قبل هم بارانی نباشد.

$۱-\beta$: احتمال اینکه امروز بارانی باشد به شرط آنکه روز قبل هم بارانی باشد.

برای محاسبه طول نوبت بارندگی مورد انتظار به طول r در یک دورهٔ m روزه با توجه به وابستگی مرتبهٔ اول مراحل زیر را انجام می دهیم :

الف) اگر r روز اول یا r روز آخر بارندگی باشد ، داریم:

$$\alpha(۱-\beta)^{r-۱}\beta + (۱-\alpha)\,\alpha(۱-\beta)^{r-۱} \qquad (۱-۴-۳)$$

ب) اگر شروع یا ختم روزهای بارندگی غیر از روز اول یا روز آخر باشد، داریم:

$$(m-r-1)\,\alpha\,\beta\,(1-\alpha)(1-\beta)^{r-1} \qquad (۲-۴-۳)$$

با توجه به روابط (۳-۴-۱) و (۳-۴-۲) طول نوبت بارندگی مورد انتظار به طول r از رابطه زیر بدست می‌آید:

$$f(r)=\begin{cases}[\alpha(۱-\alpha)(۱+m\beta)+\alpha^{۲}\beta](۱-\beta)^{r-۱}-\alpha\beta(۱-\alpha)\,r(۱-\beta)^{r-۱} & ,(r=۱,۲,...,m-۱)\\ \alpha(۱-\beta)^{r-۱} & ,\ (r=m)\end{cases}$$

(۳-۴-۳)

اگر در رابطهٔ (۳-۴-۳)، $\alpha+\beta=۱$ باشد ، ماتریس تغییر وضعیت P برابر است با :

$$P=\begin{bmatrix}\beta & \alpha\\ \beta & \alpha\end{bmatrix}$$

و برای $j_{۱}, j_{۲}, j_{۳},..., j_n \in S$

$$\alpha=P[X_{n+۱}=۱\mid X_n=j_n,...,X_۱=j_۱]=P[X_{n+۱}=۱\mid X_n=j_n]$$

به عبارتی ، احتمال شرطی X_{n+1} به شرط $X_۱, X_۲,..., X_n$ برابر است با احتمال شرطی $X_{n+۱}$ به شرط X_n. برای $X_n=۰$ و $X_{n+۱}=۱$. داریم :

$$P[X_{n+۱}=۱\mid X_n=۰]=\alpha$$

در این حالت مدل مارکوف ککران به مدل ککران تبدیل می شود و برای $\alpha=\beta=\frac{۱}{۲}$ مدل ککران به مدل گلد تبدیل می شود.

۳-۵ مثال عملی و کاربردی برای مدل های کگران و مارکوف کگران

برای برازش دو مدل طول نوبت بارندگی ، از داده های میزان بارندگی در ایستگاه هواشناسی عثمان آباد ، طی یک دوره **۱۰** ساله استفاده شده است و طبق تعریف هر روزی که میزان بارندگی بیشتر ، یا مساوی ۲/۵ میلی متر بوده کد ۱ و در غیر اینصورت کد **۰** نسبت داده شده است [۹] و [۴۵]. کدهای ۱ کنار هم طول نوبت بارندگی مشاهدات را تشکیل داده و طول نوبت بارندگی مورد انتظار برای مدل کگران از رابطه (۳-۳-۳) و برای مدل مارکوف کگران از رابطه (۳-۴-۳) پس از برآورد پارامتر(ها) بدست آمده است. لازم به ذکر است که پارامتر مدل اول با استفاده از روش درستنمایی ماکزیمم[1] و پارامتر های مدل دوم با استفاده از روش کمترین مربعات خطا[2] ، برآورد شده اند .

برای مقایسه دو مدل کگران و مارکف کگران ، طول نوبت بارندگی مشاهده شده برای مدت **۱۰** سال محاسبه شده است و طول نوبت بارندگی مورد انتظار برای مدل های مذکور پس از برآورد پارامترها از روابط (۳-۳-۳) و (۳-۴-۳) بدست آمده که خلاصه این نتایج در جدول(۳. ۵. ۱) آورده شده است.که r طول نوبت بارندگی است . برای $r = 1$ مقدار O_1 برابر با ۱۹۲ است . یعنی طی ۱۰ سال در فصل بارندگی ۱۹۲ مورد بارندگی یک روزه بوده و روزهای قبل و بعد آن خشک بوده است. و E_{1r} و E_{2r} و E_{3r} به ترتیب طول نوبت بارندگی تحت مدل های گلد ، کگران ، و مارکوف کگران می باشد .

---

[1]-**Maximum Likelihood**
[2]-**Least Square**