



دانشگاه علوم پزشکی کرمان

دانشکده بهداشت

پایان نامه مقطع کارشناسی ارشد آمار زیستی

عنوان :

مقایسه مدل‌های سری‌های زمانی و شبکه‌های عصبی مصنوعی در پیش‌بینی مرگ و میر بیماران

مبتلا به سرطان پستان

توسط : محمد مقدسی امیری

استاد راهنما : دکتر عباس بهرامپور

سال تحصیلی : 1392-1393



# **Comparison of Time Series and Artificial Neural Networks Forecasting in Mortality of Breast Cancer**

A Thesis  
Presented to  
The Graduate Studies

By  
**Mohammad Moqaddasi Amiri**

In Partial Fulfillment  
Of the requirements for the Degree  
Master of Science in:  
**Biostatistics**

**Kerman University of Medical Sciences**

**February 2014**

## چکیده:

مقدمه و هدف: یکی از مدل‌های رایج در سری‌های زمانی مدل خود برگشت میانگین متحرک تلفیق شده (ARIMA) است. ARIMA مدل‌سازی را فقط به صورت خطی انجام می‌دهد. در دنیای واقعی، بسیاری از رفتارها غیر خطی هستند. شبکه‌های عصبی مصنوعی (ANNs) روش‌های جدیدی هستند که در پیش‌بینی سری‌های زمانی استفاده می‌شوند. این مدل‌ها هیچ پیش‌فرضی ندارند و میتوانند ارتباطات غیر خطی موجود در داده‌ها را برآورد کنند. سرطان پستان دارای مرگ و میر زیادی در میان زنان می‌باشد. ممکن است مرگ و میر سرطان پستان الگویی غیر خطی داشته باشد. هدف این مطالعه برآزش هر دو روش سنتی (ARIMA) و مدرن (ANN) در مدل‌سازی مرگ و میر سرطان پستان و مقایسه دقت پیش‌بینی آنها می‌باشد.

مواد و روش‌ها: داده‌ها تعداد مرگ بر اثر سرطان پستان در 105 ماه در استان کرمان است. پس از شناسایی مدل با نمودارهای ACF و PACF، مدل‌های ARIMA مختلف برآزش داده شد و بهترین آنها بر اساس مینیمم MSE، AIC، BIC و AICc انتخاب گردید. سپس شبکه‌های مختلف با تعداد متفاوت نورون در لایه‌های ورودی و میانی، توابع فعال‌سازی مختلف و مجموعه‌های آموزش و آزمون متفاوت برآزش داده شد و بهترین شبکه بر اساس مینیمم MSE و MAE انتخاب گشت.

یافته‌ها: بهترین مدل ARIMA بر اساس MSE،  $ARIMA(1,0,1)$  یا همان  $ARMA(1,1)$ ، و بر اساس AIC، BIC مدل  $ARIMA(0,1,1)$  بودند. ما شبکه‌های  $ANN(12,6,1)$  با 84 و 95 ماه در مجموعه آموزش را بتوان بهترین شبکه‌های ANNs انتخاب کردیم. مدل‌های ARIMA، MSE و MAE بزرگتری از بهترین مدل‌های شبکه‌های عصبی مصنوعی داشتند.

نتیجه‌گیری: در مطالعه ما عملکرد مدل‌های ANNs در برآورد پارامترها و پیش‌بینی بهتر از مدل‌های ARIMA است و اختلاف زیادی در MSE و MAE بین دو روش مشاهده گردید.

واژگان کلیدی: سری‌های زمانی، شبکه‌های عصبی، سرطان پستان، ARIMA، مرگ و میر

## Abstract

**Background & Objective:** One of the common used models in time series is Auto Regressive Integrated Moving Average (ARIMA) model. ARIMA will do modeling only linearly. In the real world, many of behaviors are nonlinear. Artificial Neural Networks (ANNs) are modern methods that be used for time series forecasting. These models can identify nonlinear relationships among data. The breast cancer has the most mortality of cancers among women. It is possible that the breast cancer mortality pattern be nonlinear. The aim of this study is, fitting the both ARIMA and ANNs models on the breast cancer mortality, and comparing the accuracy of those in forecasting.

**Methods:** The data is the number of deaths caused by breast cancer in 105 months in Kerman province. The ARIMA models are fitted and chose the best one with some diagnostic criteria. Then neural networks are constructed and the best was selected with the minimum of Mean squared Error (MSE) and root Mean Absolute Error (MAE).

**Results:** The best ARIMA model based on MSE was ARIMA(1,0,1), and based on AIC, AICc, and BIC was ARIMA(0,1,1). We chose the ANN(12, 6, 1) network with 84 and 95 months in training set as the best models of ANNs. the ARIMA models have larger MSE and MAE in forecasting than the best ANNs model.

**Conclusion:** The performance of ANNs models in parameter estimating and forecasting is better than ARIMA model, and there are large differences in MAE and MSE between these two methods.

**Keywords:** Time series, Artifitial Neural Networks, Breast cancer, ARIMA, Mortality