

Farklı Sıcaklıkların *Scymnus subvillosus*'un Bıraktığı Yumurta Sayıları Üzerine Etkilerinin Karışımli Poisson Regresyon ile Analiz Edilmesi

Abdullah YEŞİLOVA⁽¹⁾

Remzi ATLIHAN⁽²⁾

Özet: Poisson regresyonu sayıma dayalı olarak elde edilen bağımlı değişkenin analizinde kullanılmaktadır. Poisson regresyon analizi sonucunda, verilerde aşırı yayılım saptandığında alternatif analiz yöntemlerinin kullanılması gerekmektedir. Bu çalışmada, söz konusu alternatif yöntemlerden biri olan karışımli Poisson regresyonu incelenmiştir. Karışımli model yaklaşımı, veri kümesinin heterojen bir yapı gösterdiğini varsaymakta ve bu heterojenliği popülasyonu kendi içerisinde alt popülasyonlara bölerek gidermektedir. Çalışmada, *Scymnus subvillosus* dişilerinin ovipozisyon süresince bıraktıkları yumurta sayıları altı alt popülasyona ayrılmıştır. Uygun alt popülasyon sayısını belirlemede AIC ve BIC ölçütleri kullanılmıştır. Parametre tahminleri EM yaklaşımı kullanılarak ML yöntemi ile elde edilmiştir. S. subvillosus dişilerinin ovipozisyon süresince bıraktıkları yumurta sayılarına ilişkin en yüksek değer 25°C de elde edilmiş olup onu, 20 ve 30°C'lerde elde edilen değerler izlemiştir. 35°C de elde edilen değer denemede kullanılan diğer üç sıcaklıkta elde edilenden oldukça düşük bulunmuştur.

Anahtar kelimeler: Aşırı yayılım, Poisson regresyonu, karışımli Poisson regresyon

Analysing the Effects of Different Temperatures on Egg Numbers of *Scymnus subvillosus* Using Mixture Poisson Regression

Abstract: Poisson regression is used in the analysis of dependent variables obtained by counting. Alternative analysis methods are necessary when there is overdispersion in data as a result of Poisson regression. In the present study, one of these alternative methods, mixture Poisson regression was studied. Mixed model approaches assume when data sets is heterogeneous, and it removes the heterogeneity by dividing the population into the sub-populations. In the study, the eggs laid by *Scymnus subvillosus* females during oviposition were divided into six sub-populations. AIC and BIC criteria were used to the determine the suitable sub-population numbers. Parameters estimations were obtained with ML method by using EM approach. The highest egg numbers was obtained at 25°C and followed by the values obtained at 20°C and 30°C. The value obtained at 35°C was found much lower than those obtained in other three temperatures.

Key words: Overdispersion, Poisson regression, mixture Poisson regression

Giriş

Sayıma dayalı olarak elde edilen verilerin analizinde Poisson regresyonu yaygın olarak kullanılmaktadır. Poisson dağılımının en belirgin özelliği, ortalama ile varyansın birbirine eşit olmasıdır. Ancak çoğu uygulamada, bu eşitliği sağlamak mümkün olmamaktadır. Poisson dağılımında söz konusu eşitlik sağlanmadığı durumda, ya aşırı yayılım ya da az yayılım görülmektedir. Poisson dağılımında, varyansın ortalamadan büyük olması aşırı yayılım, varyansın ortalamadan küçük olması ise az yayılım olarak bilinmektedir. Böyle durumlarda, veri kümelerine bilinen Poisson regresyonu uygulanmamaktadır. Bunun yerine aşırı yayılımı açıklayan yayılım parametresini (dispersion parameter) içeren regresyon modelleri kullanılmalıdır (McCullagh ve Nelder., 1989; Breslow., 1990; Dean., 1992; Wang ve ark., 1996). İstatistiksel kaynaklarda, Poisson regresyonunda oluşan aşırı yayılımı açıklayan iki yaklaşımdan bahsedilebilir. Bunlardan ilki, quasi olabilirlik

yaklaşımı, diğer yaklaşım ise karışımli Poisson model yaklaşımıdır (Wang ve ark., 1998).

Karışımli Poisson regresyon (Mixture Poisson regression= MPR), veri setinde aşırı yayılım olduğu durumlarda kullanılır. MPR, aşırı yayılıma genellikle gözlenemeyen heterojenliğin (latent heterogeneity) neden olduğunu varsaymaktadır (Dempster ve ark., 1977; Wang ve Putterman., 1998; Jensen, 1993). Örnek, farklı alt popülasyonlardan oluşan bir popülasyondan elde edilmiş olabilir. Başka bir ifadeyle, örneğin, tek bir popülasyondan değil birden fazla popülasyondan elde edilmiş heterojen bir veri kümesi olduğu varsayılır. MPR'de amaç, bir örnekteki gözlemlerin popülasyondaki gözlenmemiş alt popülasyonlara (sub-populations) ait olabileceği ve bu alt popülasyonların belirlenmesidir (Wang ve ark., 1996; Yeşilova, 2003). Veri kümesinin sınırlı sayıda alt popülasyona sahip olduğu varsayılarak sonlu karışımli

⁽¹⁾ Yazışma Adresi: Yüzüncü Yıl Üniversitesi, Ziraat Fakültesi, Zootekni Bölümü, 65080-VAN, yesilova@yyu.edu.tr

⁽²⁾ Yüzüncü Yıl Üniversitesi, Ziraat Fakültesi, Bitki koruma Bölümü, 65080-VAN

modeller (Finite mixture model) kullanılır. MPR'de parametre tahminleri, EM (Expectation Maximization) algoritması kullanılarak en çok olabilirlik (Maksimum Likelihood=ML) yöntemi ile elde edilir. Veri kümesini en iyi açıklayan modelin seçiminde, Akaiki'nin bilgi ölçütü (Akaiki Information Criteria=AIC) ve Bayesian bilgi ölçütü (Bayesian Information Criteria=BIC) en çok kullanılan uyum ölçütleridir (Dalrymple ve ark., 2002; SAS, 2006).

Bu çalışmada, karışımli Poisson Regresyon modelinin teorik özellikleri incelenerek, Bitki koruma alanında elde edilen gerçek bir kümesine uygulanmıştır. İlk olarak karışımli Poisson regresyon modelin, veri kümesinin tek bir popülasyondan elde edilmiş şekilde uyumu yapılarak, veri kümesindeki heterojenliği saptanmıştır. Daha sonrasında, söz konusu heterojenliği gidermek için elde edilen alt popülasyonlara göre modelin uyumu yapılmış, her alt popülasyon için parametre tahmini ve alt popülasyona düşen bireylerin oranları elde edilmiştir.

Materyal ve Yöntem

Materyal

Çalışmanın materyalini av olarak doğada erik ve kayısı plantasyonları ile sazlık alanlardan elde edilen *H. pruni*, avcı olarak da kayısı plantasyonlarından toplanan *S. subvillosus* oluşturmuştur. Doğadan toplanan *Scymnus subvillosus* (*S. subvillosus*) bireylerinden laboratuarda yeni bir döl elde edilmiş ve denemelere bu dölün bireyleri ile başlanmıştır. Laboratuarda üretilen dölün erginleri tarafından bırakılan yumurtalar, tabanında ısıtılmış kurutma kağıdı bulunan 5.5x1.4 cm ebadındaki petrilere tek tek alınmıştır. Yumurtalardan çıkan larvalar buldukları petri kabı içinde *H. pruni* ile pupa oluncaya kadar beslenmişlerdir. Pupadan çıkan erginler bir dişi ve bir erkek birlikte olmak üzere petri kaplarına alınmış ve ölünceye kadar tüketebileceğinden fazla besin verilmiştir. *S. subvillosus* dişilerinin farklı sıcaklıklarda ovipozisyon süresince bıraktıkları yumurta sayıları günlük gözlemlerle belirlenmiştir. Çalışma 20, 25, 30 ve 35±1 °C sıcaklık, %60±10 orantılı nem ve 16 saat aydınlatmalı iklim dolabında yürütülmüştür.

Yöntem

Yöntem üç başlık altında incelenmiştir. İlk olarak karışımli Poisson regresyon, sonrasında ise parametre tahmini ve son aşamada ise uyum ölçütleri aşamaları verilmiştir. *S. subvillosus* dişilerinin ovipozisyon süresince bıraktıkları yumurta sayıları bağımlı değişken, farklı sıcaklıklar ise bağımsız değişken olarak modele alınmıştır. Çalışmada, gerekli analizler SAS (2006) istatistiksel yazılım programı kullanılarak yapılmıştır.

Karışımli Poisson regresyon

Poisson regresyonunda model,
 $E(Y_i/X_i) = \lambda_i = \exp(X_i\beta)$ $i=1, 2$,

biçiminde yazılmaktadır. Eşitlikte, Poisson ortalaması (λ_i) ile ortak değişkenler arasındaki ilişki bir bağlantı (link) fonksiyonu ile verilmiştir. Veri kümesi aşırı yayılım gösterdiği zaman Poisson regresyonun kullanılması doğru olmayan parametre tahminlerine neden olmaktadır. Karışımli model için kesikli karışım dağılışı,

$$p(y) = \sum_{k=1}^K P(y/v_k \exp(\beta'x)) \pi_k \quad 1$$

biçiminde yazılabilir. 1 numaralı eşitlikte, β regresyon katsayılarına ait vektör, v gamma dağılımlı tesadüfi etki veya değişken, y_i gözlem değerleri $v_k \exp(\beta'x)$ ortalamalı Poisson dağılımından elde edilmektedir. K alt popülasyon sayısı ve π_k , k'nci alt popülasyonun karışma olasılığıdır (Wang ve ark., 1996; Wang ve ark., 1998; Yeşilova, 2003).

$y_i = \{y_1, \dots, y_n\}$ Poisson dağılımına sahip gözlem değerlerinden oluşan veri kümesinin homojen tek bir popülasyonu temsil etmeyip birden fazla alt popülasyona ait heterojen bir veri kümesi olabilir. Başka bir ifade ile veri kümesinde gözlenemeyen alt popülasyonlar bulunabilir. Bu durumda y değerlerine ilişkin marjinal olasılık fonksiyonu,

$$f(y) = \sum_{k=1}^K P(C=k)P(Y=y|C=k) = \sum_{k=1}^K \pi_k f(y, \lambda_k) \quad 2$$

şeklinde yazılabilir. Poisson dağılışı veri setinin, K kadar alt popülasyona ait heterojen bir örnek olması durumunda k'nınci alt popülasyona giren i'inci şans değişkeninin olasılığı,

$$\pi_{ik} = P(c_i = k)$$

biçiminde verilebilir. Bu durumda,

$$\sum_{i=1}^K \pi_{ik} = 1$$

olmaktadır. Bütün veriler için log-olabilirlik fonksiyonu,

$$\sum_{i=1}^n \sum_{k=1}^K c_{ik} \log \pi_k + \sum_{i=1}^n \sum_{k=1}^K c_{ik} \log P(y_i/\lambda_i) \quad 3$$

biçiminde verilebilir (Wang ve ark., 1996; Wang ve ark., 1998; Dempster ve ark., 1977). 3 nolu eşitlikte, C gözlenemeyen gözlemler (alt popülasyon sayısı) olup,

$$C = \{c_{ik}, i = 1, 2, \dots, n; k = 1, 2, \dots, K\}$$

Burada C_{ik} ,

$$\left\{ \begin{array}{l} C_{ik} = 1, C_{ik} \in K \\ C_{ik} = 0, \text{diger durumlarda} \end{array} \right\}$$

olarak verilebilir. Karışımı Poisson regresyonun ortalaması ve varyansı,

$$E(y_i) = E(E(y_i | K = k)) = \sum_{k=1}^K \pi_k \lambda_k \quad 4$$

$$\begin{aligned} Var(Y_i) &= E\{Var(Y_i | K = k)\} + Var\{E(Y_i | K = k)\} \quad 5 \\ &= \sum_{k=1}^K \pi_k \lambda_k \left\{ \sum_{k=1}^K \pi_k \lambda_k^2 - \left(\sum_{k=1}^K \pi_k \lambda_k \right)^2 \right\} \end{aligned}$$

biçiminde verilebilir. 4 ve 5 numaralı eşitliklerde varyans ortalamadan büyük olduğundan dolayı aşırı yayılım söz konusu olur. Veri setinin heterojen olmadığı veya aşırı yayılım göstermediği durumlarda, ortalama ile varyans arasındaki ilişki,

$$E(Y_i) = Var(Y_i)$$

ve yukarıda verilen 5 numaralı eşitlikteki varyans formülünde,

$$Var(y_i / K = k) = 0$$

olur. Böylece ortalama ile varyans arasındaki eşitlik sağlanmış olur.

ML yöntemi ve EM algoritması

Karışımı Poisson regresyonunda, parametre tahminleri E ve M aşamalarında oluşan EM algoritması kullanarak en çok olabilirlik yöntemi ile elde edilmektedir. E aşamasında, gözlenmiş verilerin koşullu beklenen değerleri üzerinden eksik gözlemleri türetilir. Burada eksik gözlemler alt populasyon sayısıdır. Başlangıçta veri kümesinin kaç alt populasyona ayrılacağı bilinmediğinden dolayı, E aşamasında söz konusu alt populasyon sayısı belirlenir. M aşamasında ise tüm veriler için beklenen log-olabilirliği maksimize eden parametreler elde edilir (Dempster ve ark., 1977; Jansen, 1993; Dalrymple ve ark., 2002). Model için EM algoritmasının aşamaları aşağıdaki gibi verilebilir;

Birinci aşamada, $\beta^{(0)}$ ve $\pi_k^{(0)}$ başlangıç değerleri \mathcal{E} ve \mathcal{E}_0 tolerans ile belirlenir.

E aşamasında, $\beta^{(0)}$ ve $\pi^{(0)}$ başlangıç değerleri verildiğinde gözlenmiş veriler (X, Y) ile parametrelerin

başlangıç değerleri üzerinden, C eksik gözlemleri elde edilir. $\hat{C}_{ik}(\beta^{(0)}, \pi^{(0)})$ kullanılarak C_i 'nin k'nıncı unsurunun koşullu olasılığı,

$$\hat{c}_{i,k}(\beta^{(0)}, \pi_k^{(0)}) = \frac{\pi_k f_k(y_i/x_i, \beta_k^{(0)})}{\sum_{k=1}^K \pi_k f_k(y_i/x_i, \pi_k^{(0)})}, k=1,2,\dots,K \quad 6$$

biçiminde verilebilir.

M aşamasında,

$$\left\{ c_i(\beta^{(0)}, \pi_k^{(0)}) = (z_{i,1}, \dots, z_{i,K})' ; i = 1, 2, \dots, n \right\} \quad \text{koşullu}$$

olasılıkları verilmişken, parametre tahminleri log olabilirlik fonksiyonun β ve π 'ya göre maksimize edilmesi ile,

$$Q = (Q_1, Q_2) = E\{L(Y, C, \beta, \pi, X) | Y, X, \beta^{(0)}, \pi_k^{(0)}\} \quad 7$$

$$Q = Q_1 + Q_2$$

ve buradan, Q_1 ve Q_2 ,

$$Q_1 = \sum_{i=1}^n \sum_{k=1}^K c_{i,k}(\beta^{(0)}, \pi_k^{(0)}) \log(\pi_k) \quad 8$$

$$Q_2 = \sum_{i=1}^n \sum_{k=1}^K c_{i,k}(\beta^{(0)}, \pi_k^{(0)}) \log(y_i/\lambda_k) \quad 9$$

elde edilir. Burada $\hat{\beta}$ ve $\hat{\pi}$ tahmin edicileri, Q_1 ve Q_2 eşitliklerinin π ve β 'ya göre türevlerinin alınması ile,

$$\frac{\partial Q_1}{\partial \pi_k} = \sum_{i=1}^n \left(\frac{\hat{c}_{i,k}}{\hat{\pi}_k} - \frac{\hat{c}_{i,K}}{\hat{\pi}_K} \right) = 0, k = 1, \dots, K-1 \quad 10$$

$$\frac{\partial Q_2}{\partial \beta} = \sum_{i=1}^n \sum_{k=1}^K \hat{c}_{i,k} \frac{\partial}{\partial \beta} P(y_i/\lambda_k) = 0 \quad 11$$

biçiminde elde edilir. 10 numaralı eşitlik kullanılarak $\hat{\pi}_k$,

$$\hat{\pi}_k = \frac{1}{n} \sum_{i=1}^n \hat{c}_{i,k}, k=1, \dots, K-1$$

biçiminde elde edilmektedir (Wang ve Putterman., 1998; Wang ve ark., 1996; Wang ve ark., 1998).

Model uyumu

AIC ve BIC karışımı model için yaygın olarak kullanılan uyum ölçütleridir (Wang ve Putterman, 1998; Dalrymple ve ark., 2002). Genel olarak uyum ölçütleri;

$$AIC = -2LL + 2m$$

$$BIC = -2LL + m \log(n)$$

biçiminde yazılabilir. Burada,

LL: Karışımli Poisson regresyon modelinde iterasyon bittikten sonra elde edilen log-olabilirlik değeri,

m: parametre sayısı

n: örnek büyüklüğü olmaktadır.

Bulgular

Günlük bırakılan yumurta sayıları sayıma dayalı olarak elde edildiğinden dolayı ilk olarak veri kümesine Poisson regresyonu uygulanmıştır. Poisson regresyonuna ilişkin uyum istatistikleri Çizelge 1’de verilmiştir. Poisson regresyonunda uyum istatistikleri olarak devians (sapma) ve Pearson Khi-kare kullanılmaktadır. Veri kümesine Poisson regresyonun uygulanabilmesi için devians ve Pearson Khi-kare yayılım değerlerinin 1’e eşit olması gerekmektedir. Çalışmada, yayılım değerinin 1’den büyük çıkması sonucunda Poisson dağılımındaki ortalama ile varyans arasındaki eşitliğinin sağlanamadığı görülmüş ve veri kümesinde aşırı yayılım olduğu saptanmıştır.

Çizelge 1. Poisson regresyonuna ilişkin uyum istatistikleri
Table 1. Goodness of fit statistics for Poisson regression

Uyum istatistikleri Goodness of statistics	Serbestlik derecesi Degrees of freedom	Hesap değeri Value	Yayılım Dispersion
Sapma Deviance	51	1872.089	36.708
Pearson Khi-kare Pearson Chi-square	51	1896.280	37.185

Poisson regresyonunda yayılım bir değerinden büyük çıktığı için veri kümesine karışımli Poisson regresyon analizi uygulanmıştır. Uygun alt populasyon sayısının belirlenmesini sağlayan uyum ölçütleri Çizelge 2’de verilmiştir. Genel olarak en küçük AIC ve BIC uyum ölçütlerine sahip alt populasyonlu model, verilerin en iyi dağıldığı model olarak kabul edilir. Altı alt populasyonlu modele kadar uyum ölçütlerinin küçülmeye başladığı ve daha sonra ise büyüdüğü saptanmıştır. AIC ve BIC uyum ölçütleri en küçük değerlerini 6 alt populasyonlu modelde elde etmişlerdir. Bu nedenle altı alt populasyonlu model verilerin en iyi dağıldığı modeldir. 6 alt populasyonlu modelden sonraki alt populasyonların uyum ölçütleri giderek büyüdüğünden dolayı 8 alt populasyonlu modelden sonraki alt populasyonlara yer verilmemiştir. Bunun sonucu olarak altı alt populasyon veri kümesinin en iyi dağıldığı populasyon olarak seçilmiştir.

Çizelge 2. Farklı alt populasyonlar için model seçimi
Table 2. Model selection for different sub-populations

Alt populasyonlar sub-populations	Model seçimi /model selection	
	BIC	AIC
1	-1690.484	-1688.477
2	-691.261	-686.242
3	-472.531	-464.501
4	-389.569	-378.526
5	-371.429	-357.378
6	-347.243	-330.180
7	-353.046	-332.973
8	-358.540	-335.456

Karışımli Poisson regresyona göre 55 *S. subvillosus* dişisinin ovipozisyon süresince bıraktıkları yumurta sayıları tek bir populasyona ait olmadıkları ve altı alt populasyonu temsil ettikleri görülmektedir. Seçilen altı alt populasyonlu modele ilişkin parametre tahminleri ve karışma olasılıkları Çizelge 3’de verilmiştir. 55 gözlem değerinin (dişilerin) %10.91’i (6 gözlem) birinci alt populasyona, %16.36’sı (9 gözlem) ikinci alt populasyona, %23.64’ü (13 gözlem) üçüncü alt populasyona, %16.36’sı (9 gözlem) dördüncü alt populasyona, %20’si (11 gözlem) beşinci alt populasyona ve %12.73’ü (7 gözlem) altıncı alt populasyona dahil olunmuştur. Karışımli Poisson regresyonda, oluşturulan her bir alt populasyon için parametre tahmininde Poisson regresyonunu esas alınır. Bu bağlamda altı alt populasyonda da, $(Y|C=k) \sim \log(\lambda) = X'\beta$ modeli geçerlidir. Her bir alt populasyonda parametre tahminleri aşağıdaki gibi olmuştur;

$$\text{Log}(\lambda_1|C=1) = 4.693 - 0.474 * \text{sıcaklık}$$

$$\text{Log}(\lambda_2|C=2) = 5.256 - 0.395 * \text{sıcaklık}$$

$$\text{Log}(\lambda_3|C=3) = 5.338 - 0.201 * \text{sıcaklık}$$

$$\text{Log}(\lambda_4|C=4) = 5.482 + 0.048 * \text{sıcaklık}$$

$$\text{Log}(\lambda_5|C=5) = 5.214 + 0.028 * \text{sıcaklık}$$

$$\text{Log}(\lambda_6|C=6) = 5.823 + 0.059 * \text{sıcaklık}$$

Eşitliklerde de görüldüğü gibi sıcaklığın her bir alt populasyonda parametre tahminleri birbirlerinden oldukça farklı bulunmuştur. Örneğin birinci alt populasyonda sıcaklık parametre değeri -0.474 iken altıncı alt populasyonda 0.059 olarak bulunmuştur. Bu sonuçlar, *S. subvillosus* dişilerinin ovipozisyon süresince bıraktıkları yumurta sayılarının Poisson dağılımı gösterdiği, ancak tek bir populasyonu temsil etmedikleri, kendi içerisinde altı alt homojen populasyona sahip oldukları ve her alt populasyon için ayrı parametre tahminleri yapılması gerektiğini göstermektedir. Birinci, ikinci ve üçüncü alt populasyonda sıcaklığın etkisi önemli bulunmuşken ($P > 0.05$) dördüncü, beşinci ve altıncı alt populasyonda önemsiz bulunmuştur ($P < 0.05$).

Çizelge 3. Alt populasyonlar için parametre tahminleri
Table 3. Parameter estimation for sub-populations

Alt populasyonlar Sub-populations	Karışma olasılıkları Mixing probability	Parametre tahmini Parameter estimation	
		Intercept Tahmin (std. Hata) Estimation (std. Error)	Sıcaklık /temperature Tahmin (std. Hata) Estimation (std. error)
1	%10.91	4.693 (0.197)**	-0.474 (0.065)**
2	%16.36	5.256 (0.112)**	-0.395 (0.044)**
3	%23.64	5.338 (0.087)**	-0.201 (0.032)**
4	%16.36	5.482 (0.066)**	0.048 (0.029)
5	%20.00	5.214 (0.113)**	0.028 (0.044)
6	%12.73	5.823 (0.062)**	0.059 (0.033)

** P<0.01

Çizelge 4'te farklı sıcaklık değerlerine göre 55 *S. subvillosus* dişisinin ovipozisyon süresince bıraktıkları toplam yumurta sayılarının alt populasyonlara göre dağılımları verilmiştir. En düşük toplam yumurta sayısı 35 °C de elde edilmiş olup, Çizelge 4.'te de görüldüğü gibi bu sıcaklıktaki bireyler 1, 2 ve 3. alt populasyonu dahil olmuşlardır. Yumurta sayısı bakımında birbirine oldukça benzerlik gösteren 20 °C ve 25 °C deki bireyler ise en yüksek toplam yumurta sayısının yer aldığı 4 ve 6. alt populasyonlarda yer almışlardır. Gözlem değerlerinin (yumurta sayılarının) alt populasyonlarına göre dağılımları Şekil 1'de verilmiştir. Şekil 1'de bırakılan yumurta sayının en çok altı alt populasyonlu modelde, daha sonra dört, beş, üç, iki ve en az ise bir alt populasyonlu modellerde yer almaktadır. Bırakılan toplam yumurta sayısı bakımından alt populasyonlar arasındaki farklılık Şekil 1 üzerinden daha iyi görülmektedir. Elde edilen sonuçlar bakımından Çizelge 4 ile Şekil 1 benzerlik göstermektedir. Karışımli Poisson regresyonu kullanılarak, altı alt populasyonlu modelin en uygun model olduğu AIC ve BIC uyum ölçütlerine bakılarak belirlenmiştir. Daha sonra seçilen altı alt populasyonlu modelin her alt populasyonu için ayrı parametre tahminleri elde edilmiştir.

Çizelge 4. Farklı sıcaklık değerlerine göre 55 gözlem değerinin alt populasyonlara göre dağılımı

Table 4. Distribution of 55 observation values based on the different temperature degrees and sub-populations

Alt populasyonlar Sub-populations	Sıcaklık dereceleri temperature degrees			
	20°C	25°C	30°C	35°C
Alt populasyon 1	1	0	0	5
Alt populasyon 2	2	2	1	4
Alt populasyon 3	4	2	4	3
Alt populasyon 4	3	3	3	0
Alt populasyon 5	1	4	6	0
Alt populasyon 6	3	3	1	0
Bırakılan toplam yumurta Egg numbers	216.43	224.93	200.40	40.08

Şekil 1'de görüldüğü gibi veri kümesi altı alt populasyona ayrılarak söz konusu heterojenlik giderilmiştir.

Böylece alt populasyonlar içi homojenlik sağlanırken alt populasyonlar arası heterojenlik ortaya konmuştur.

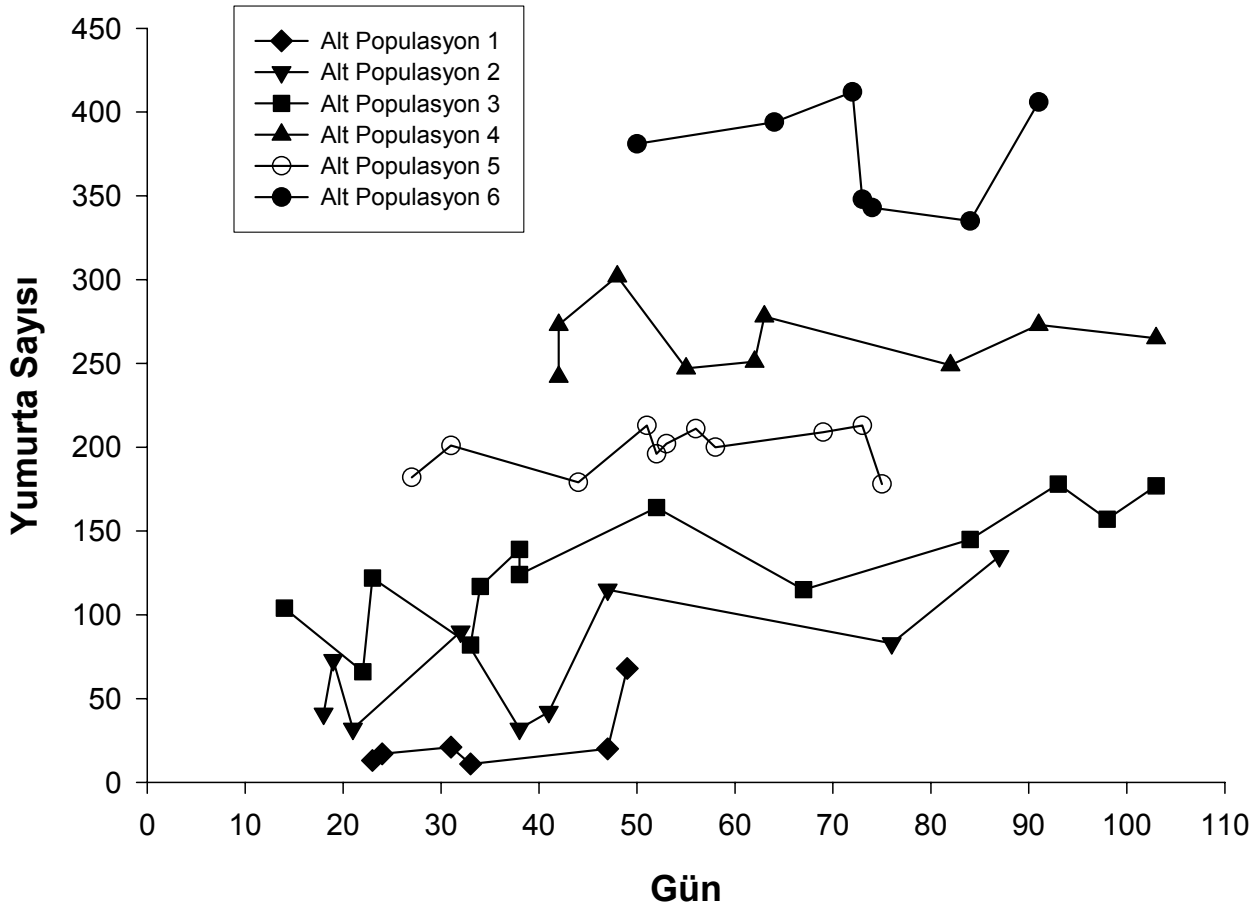
Tartışma ve Sonuç

Poisson regresyonunda model uyumu devians ve Pearson Khi-kare uyum istatistikleri kullanılarak yapılmaktadır. Her iki uyum istatistiğine ilişkin yayılım değerlerinin 1'e eşit olması yayılım olmadığı, 1'den büyük olması aşırı yayılım ve 1'den küçük çıkması da az yayılım olarak tanımlanmaktadır. Çalışmalarda, genellikle aşırı yayılım gözlenmektedir (Wang ve ark., 1996; Wang ve ark., 1998). Poisson regresyonu sonucunda uyum istatistikleri için yayılım değerinin, 1'den oldukça büyük çıkması aşırı yayılıma neden olmuştur. Böyle durumlarda aşırı yayılımı dikkate alan regresyon modellerinin kullanılması gerekmektedir. Bu çalışmada aşırı yayılım parametresini dikkate alan karışımli Poisson regresyon uygulanmıştır.

Karışımli Poisson regresyon analizi değişik alt populasyonlar (2, 3, 4, 5, 6, 7, 8 alt populasyonlu modeller) için yapılmıştır. En küçük AIC ve BIC değerlerine sahip model en iyi model olarak kabul edilmektedir (Wang ve ark., 1996; Wang ve ark., 1998; Dalrymple ve ark., 2002). Bunun sonucunda, AIC ve BIC uyum ölçütleri kullanılarak altı alt populasyonlu modelin uygun olduğu tespit edilmiştir. Model uyumu yapılan alt populasyonlar içerisinde altı alt populasyona ilişkin hem AIC hem de BIC uyum ölçütlerinin diğer alt populasyonlara göre daha küçük olduğu saptanmıştır. Bundan dolayı, yalnızca altı alt populasyonlu model için parametre tahminleri verilmiştir. Zaten karışımli Poisson regresyonunda amaç, verilerin dağıldığı uygun alt populasyon sayısını belirleyip, yalnızca seçilen alt populasyonlu model için parametre tahmini yapmaktır. Çalışmada altı alt populasyonlu modelden sonraki alt populasyonlu modellerin uyum ölçütleri giderek büyüdüğünden dolayı 8 alt populasyonlu modelden sonrasına yer verilmemiştir. Yapılan çalışmalarda, genellikle seçilen alt populasyonlu modelden sonraki bir veya iki alt populasyonlu modele ilişkin uyum ölçütleri verilmektedir (Wang ve ark., 1996; Wang ve ark., 1998;

Dalyrmpile ve ark., 2002; Yeşilova, 2003). Bu nedenle seçilen altı alt populasyondan sonraki 7 ve 8 alt populasyonlu modellere ilişkin AIC ve BIC değerleri Çizelge 2’de verilmiştir. Karışımı Poisson regresyonun kullanıldığı durumlarda, en uygun modeli belirlemede söz konusu ölçütlerin birlikte değerlendirilmesi gerekmektedir.

Bazı çalışmalarda her iki uyum ölçütünün birlikte değerlendirilmesine karşın (Wang ve ark., 1996; Wang ve ark., 1998; Yeşilova, 2003), bazı çalışmalarda ise BIC uyum ölçütünün yeterli olabileceği belirtilmektedir (Dalyrmpile ve ark., 2002).



Şekil 1. Yumurta sayılarının alt populasyonlara dağılımı
Figure 1. Distribution of egg numbers to sub-populations).

Şekil 1’de farklı sıcaklık düzeyleri esas alınarak *S. subvillosus* dişilerinin ovipozisyon süresince bıraktıkları yumurta sayılarına göre alt populasyonlar verilmiştir. Şekil 1’de gözlem değerlerinin alt populasyonlara göre dağılımına bakıldığında, alt populasyonlar birden altıya doğru gidildiğinde bırakılan toplam yumurta sayılarının da arttığı gözlenmektedir. Yani en çok yumurta 6’ncı alt populasyonda daha sonra sırayla 4’üncü, 5’inci, 3’üncü, 2’inci alt populasyonda yer almaktadır. En az yumurta bırakılan alt populasyon ise 1’inci alt populasyondur. Çizelge 4’te, özellikle 35°C’nin birinci, ikinci ve üçüncü alt populasyonlar üzerinde etkili iken dördüncü, beşinci ve altıncı alt populasyonların oluşumu üzerinde etkili olmamıştır. Buna karşın 25°C ve 30°C sıcaklıkları birinci

alt populasyonun oluşumunda etkili olmamışlardır. 20°C ve 25°C sıcaklıkları bırakılan yumurta sayıları altıncı ve dördüncü alt populasyonlu modellerde daha etkili olmuşlardır. *S. subvillosus* dişilerinin ovipozisyon süresince bıraktıkları yumurta sayılarına ilişkin en yüksek değer 25 °C’de elde edilmiş olup, onu 20 ve 30 °C’lerde elde edilen değerler izlemiştir. 35 °C’de elde edilen değer denemede kullanılan diğer üç sıcaklıkta elde edilenden oldukça düşük bulunmuştur. Çizelge 4’de görüldüğü gibi sıcaklık artınca bırakılan yumurta sayıları azalmaktadır. Coccinellidlerle yapılan çeşitli araştırmalarda bu çalışmada elde edilenlere benzer sonuçlar bildirilmiştir (Uygun ve Atlıhan,, 2000, Atlıhan ve Özgökçe, 2002).

Sonuç olarak bağımlı değişkenin sayıma dayalı olarak elde edildiği durumlarda Poisson regresyonun kullanılması gerekmektedir. Ancak Poisson regresyonunda ortalama ile varyans arasındaki eşitsizlikten kaynaklanan aşırı yayılım olduğu durumlarda, alternatif yöntemlerin kullanılması gerekmektedir. Bu çalışmada veri kümesinde aşırı yayılım olduğu durumlarda karışımli Poisson regresyonun uygulanabilirliği incelenerek, bu yöneme ilişkin bir uygulama yapılmıştır.

Kaynaklar

- Atlıhan, R., Özgökçe, M.S., 2002. Development, fecundity and prey consumption of *Exochomus nigromaculatus* feeding on *hyalopterus pruni*. *Phytoparasitica*, 30(5): 443-450.
- Breslow, N., 1990. Tests of hypotheses in overdispersed poisson regression and other quasi-likelihood models. *Journal of American Statistical Association*, 85(410):565-571.
- Dalrymple., Hudson, I.L., Ford, R.P.K., 2002. *Finite Mixture, Zero-Inflated Poisson and Hurdle Models with Application to SIDS*. University of Canterbury, Christchurch, New Zealand, 19.
- Dean, C.B., 1992. Testing for overdispersion in poisson and binomial regression models. *Journal of American Statistical Association*, 87(418):451-457.
- Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM Algrithm. *Journal of Royal Statistical Society*, 39: 1-18.
- Jansen, R.C., 1993. Maksimum Likelihood in a Generalized Linear Finite Mixture Model by Using the EM Algorithm. *Biometrics*, 49(1):227-231.
- McCullagh, P., Nelder, J.A., 1989. *Generalized Linear Models*. Second Edition, Chapman and Hall, London, 486.
- SAS, 2006. *SAS/STAT Software:Hangen and Enhanced*. SAS, Inst. Inc., USA.
- Uygun, N., Atlıhan, R., 2000. The effect of temperature on development and fecundity of *Scymnus levaillanti*. *BioControl*, 45: 453-462.
- Yeşilova, A., 2003. *Biyolojik Çalışmalardan Elde Edilen Kategorik Verileri Karışık Poisson Regresyon Analizinin Uygulaması* (Doktora Tezi, Basılmamış), YYÜ Fen Bilimleri Enstitüsü, Van.
- Wang, P., Puterman, M.L., Cockburn, I.M., Le, N., 1996. Mixed poisson regression models with covariate dependent rates. *Biometrics*, 52:381-400.
- Wang, P., Cockburn, I.M., Puterman, M.L., 1998. Analysis of patent data- mixed poisson regression model approach. *Journal of Business and Economic Statistics*, 16(1):27-41.
- Wang, P., Putterman, M.L., 1998. Mixed logistic regression models. *Journal of Agriculture, Biological and Environmental Statistics*, 3(2):175-200.