

## 16. 連続型確率分布を用いたダイズの節間長の近似に関する研究

環境立地学研究室 KN

### 1. 目的と背景

近年植物の計測の精度やデータ数が向上されたことにより、植物の立体形状のモデル化が行えるようになってきた。従来の植物のモデリングの場合、従来の研究では個体差の近似には、二項分布や負の二項分布などの離散型確率分布が使用されてきた。しかし、節間長などは実数であるため、離散型確率分布で表すには適していない。そこで本研究では代表的な連続型確率分布である、ガンマ分布とベータ分布と正規分布でダイズの節間長を近似するために、ダイズを研究対象とし、節間長、節数を測定し、植物形状とその成長を再現するために必要な、節間長の節毎の変化とその分布について分析を行った。

### 2. 研究方法

使用した節間長のデータは1992年に長野県中信農業試験場で栽培されたダイズのデータを用いた。データの概要を表1に示す。またダイズの各部分の名称は図1に示したように、主軸は根元から順に $U_1, U_2, \dots, U_n$ とし、側枝も同様に主軸の付け根から数える。また節間長の近似は、各部位ごとの節間長を1mmごとの階級に分け、確率量の期待度数と累積分布について行った。そして確率分布の適合性を調べるために $\chi^2$ 検定を行った。 $\chi^2$ は式1で表される。

$$\chi^2 = \sum_{i=1}^k \frac{(n_i - m_i)^2}{m_i} = \sum_{i=1}^k \frac{(n_i - n_i p_i)^2}{n_i p_i} \quad \dots (1)$$

( $n_i$ :観測度数  $m_i$ :期待度数  $p_i$ :確率)

統計量 $\chi^2$ は標本数 $n$ が大きく、各 $m_i$ が5以上であれば近似的に自由度 $=k-1$ の $\chi^2$ 分布に従う。また有意水準 $\alpha$ %の棄却域は式2で表される。

$$\chi^2 > \chi^2_{\alpha}(k-1) \quad \dots (2) \quad (k:階級の個数)$$

### 3. 結果と考察

ある階級における節間長とその節間長が発生する期待度数との関係を、ダイズAを用いて示した例が図2である。またある階級における節間長とその節間長が発生する確率を累

表1 使用したダイズのデータ

播種時期と 栽植密度	名称と本数		
	タチナガハ	ワセシロゲ	タマホマレ
7月1日 0.05m <sup>2</sup> /本	A 65本	C 51本	E 51本
6月1日 0.11m <sup>2</sup> /本	B 53本	D 59本	F 50本

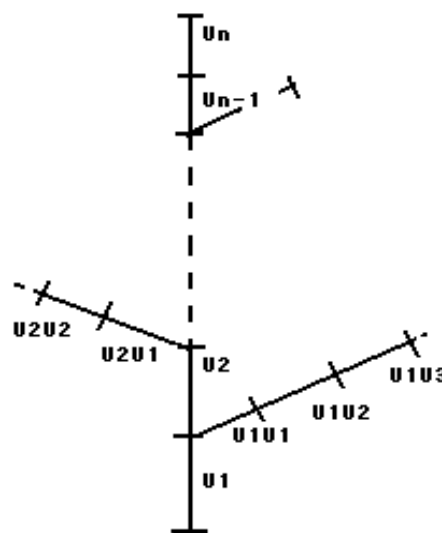


図1 各部位の名称

積していったものが図3である。これらの近似した値はどの連続型確率分布に当てはめてみても実測値とよく一致していて、この結果はA以外のサイズにも当てはまった。

また図4は節間長をベータ分布で近似し、2検定を行った結果である。有意水準5%で検定を行い、分布がその部位に当てはまるなら採択(○)、当てはまらなかったら棄却(×)、検定が行えなかったら検定不適(△)と表記した。検定が不適となるのは期待度数  $m_i$  が5以下の場合である。検定不適の部位は枝の先端部分などでサンプル数が少ないことが原因であるが、RMSE という指標を用いこの部分の誤差を調べた。RMSE は式3で表される。

$$RMSE = \sqrt{\frac{\sum (F_i - A_i)^2}{N}} \quad \dots (3) \quad (F_i : \text{理論値} \quad A_i : \text{実測値} \quad N : \text{階級数})$$

図5に枝別の分布ごとのRMSE値を算出した例を示す。図4ではサイズAのU1UNの部位は検定不適となりその精度が不明であるが、図5を見ると枝の先端部までRMSE値が低い数値で推移していることがわかる。これは検定で不適となった他の部位にもこの傾向は見られた。以上より図4で検定不適となっている部位にも、節間長の近似に連続型確率分布を用いることが有効であることがわかった。

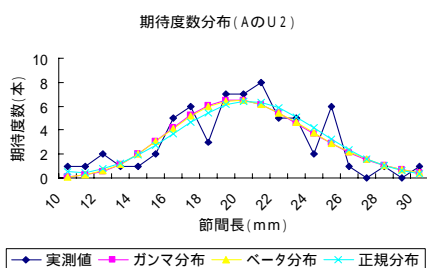


図2 サイズA(U2)の期待度数分布

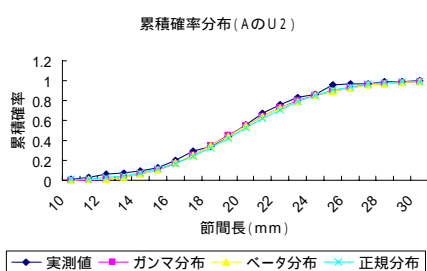


図3 サイズA(U2)の累積確率分布

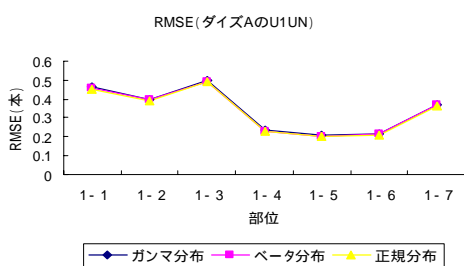


図5 サイズA(U1UN)のRMSE値

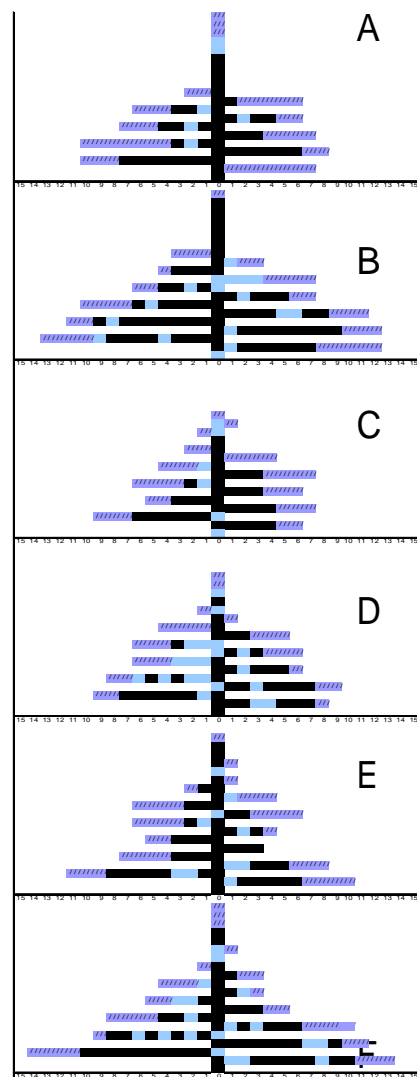


図4 ベータ分布の検定結果

