

【原著】

入学者数予測のための簡便な回帰分析法

菅田節朗（慶應義塾大学薬学部）

若干の仮定からなる単純な「歩留率モデル」を考えた。その主な要点は、1)入試の総得点を偏差値に変換した；2)偏差値と各偏差値における歩留率（点歩留率）との相関を表すモデル曲線にはロジスティック曲線を用いた、この2点である。このモデルに基づく回帰分析法により本学薬学部の2008年度および2009年度入試の回帰分析を行った。その結果を用いて2010年度入試の入学者数予測を行い、よい結果を得た。歩留率の意味についても論じた。

1 はじめに

各大学・学部・学科が入学定員充足率を適正に保つことはきわめて重要なことである。充足率が低すぎるのは、大学にとって経営上のマイナスとなる。入学した学生達にとっても、仲間が少ないということは必ずしも歓迎すべきことではないだろう。逆に充足率が高すぎるのは、教室の収容人数等の面で大学は困ることになるし、学生達にとっては教育環境の悪化ということになる。さらにその度が過ぎると、特に私立大学では私立大学等経常費補助金が配分されないという経営上のマイナスに直面することになる。

各大学、特に私立大学の平成22年度(569校)の入学定員充足率は、充足率110%台(110%以上~120%未満)の131校(569校中の23%)をピークに、充足率50%未満(13校)から充足率130%以上(17校)まで幅広く分布している(日本私立学校振興・共済事業団 私学経営情報センター、2010:3)。このような充足率分布の広がり、の主要な原因の一つは、各大学が充足率を適正に保つことの難しさであろう。つまり、入学者数予測あるいは歩留率予測の難しさということであろう(歩留率=入学者数/合格者数)。

入学者数予測では各大学は独自にさまざまに工夫しているようだが、特に私立大学はほとん

ど例外なく大変苦勞しているようである(朝日新聞、2008)。しかし、入試は一般に機密情報が多いためか、その方法はほとんど公表されていない。志願者数などから合格者数を決める経験式を出している例(福田、2003)、さらにオープンキャンパス参加者数を用いて改良した経験式を出している例(福田、2005)、併願校の情報などをもとに個々の受験生の入学確率を推定して合格者数を決める方法(小林・高野、2004)など、著者の探しえた文献はごくわずかである。

最近の数年間分の歩留率の推移を見ながら入学者数予測を行う大学もあるかと思う。著者の属する学部(前身の共立薬科大学時代を含めて)も最近まで(途中1年を除き)そのような学部であった。しかし、歩留率は受験者数や合格者数に左右される性質を有するので(5.4参照)、それのみで適正な入学者数を予測するのは困難である。著者は、一般的な意味の歩留率ではなく、入試の総得点(実際はそれを偏差値に変換)の「各点における歩留率(点歩留率と名づける)」に着目し、総得点と点歩留率の間に相関がある場合に適用できる簡便な回帰分析法を提案したい。この方法は20余年前にはほぼ完成したが、学部の理解を得るのは難しく、実際の入

試に適用するのは今回は9年ぶり2回目である。本方法は、自然と思ひ浮かぶごく簡単な方法であり、どこか別の大学でもやっていてふしぎはない。その意味で、公表に躊躇していた。しかし、少しでも一般性のある方法なら、機密事項に配慮しつつも公表すればこの分野の研究が進展し、全私立大学およびその受験生の利益となると思う。

2 本学薬学部の入試概要

本学薬学部の入学定員は2008年度から2010年度の間は、薬学科180名、薬科学科30名であり、主な募集方法は一般入試(A方式)、一般入試(B方式)、塾内進学の3つである。入学定員充足率(カッコ内は薬学科、薬科学科の順)は、2008年度(1.04, 1.43)は薬科学科の1.43が高すぎ、2009年度(0.98, 1.10)は薬学科の0.98が低すぎるなど不本意な点があったが、2010年度(1.08, 1.23)はほぼ意図した数値に近かった。薬科学科は定員が少ないため充足率はわずかの人数の読み違いで大きく変化している。

最も募集人員の多い一般入試(B方式)の2008年度～2010年度入試を回帰分析の対象とした(2008年度および2009年度入試の回帰分析結果をもとに2010年度入試の入学人数予測を行った)。この間の入試要項は変わらず、試験は両学科同一の試験問題で、3科目計300点満点で行われた。受験者数等は表1の通りである。受験者の平均点等は省略する。

表1 一般入試(B方式)概要 ()内は募集人員

	年度	受験者数	合格者数(補欠繰上含)
薬学科	2008	1881	411
	2009	2002	384
	(140) 2010	1810	441
薬科学科(15)	2008	531	109
	2009	401	100
	2010	402	111

3 歩留率モデル

歩留率に影響を与える因子は多様であろう。今回の回帰分析は、そのような個々の因子による分析ではなく、(そのような因子の総合された)統計結果のみで行うものである。単純な回帰分析ではあるが、その回帰分析では、いくらかの仮定を置くことでできたモデルを使用した。そのモデルをここでは「歩留率モデル」と称することとする。歩留率モデルで置いた仮定は以下の通りである(イメージを図1に示す)。

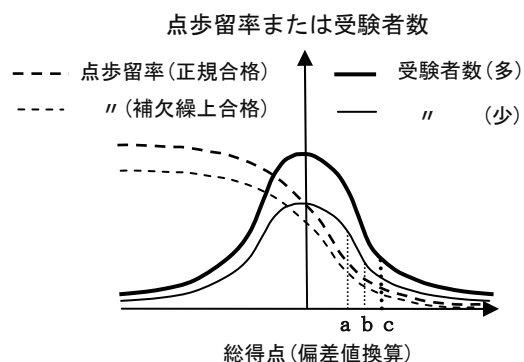


図1 歩留率モデル

- 1) 受験生の成績(総得点)の分布は正規分布となる。
- 2) 受験生全体の質は、年度により変わることなく、一定である。
- 3) したがって、受験生全体の平均点(Mean)および標準偏差(SD)の年度間の違いは入試問題の難易度の違いによるものである。
- 4) 年度間の入試問題の難易度は、線形等化法を用いた偏差値化により補正できるものとする(橋本, 2007)。したがって、総得点を次式

$$\text{偏差値} = 10 \times (\text{総得点} - \text{Mean}) / \text{SD} + 50 \quad (1)$$

で偏差値換算した受験生の成績分布は、異なる2年度分で表すと、平均点(50)と標準偏差(10)は同じで、一般に受験者数は

異なるので、図1のようになる。

- 5) 個々の総得点（実際はその偏差値）における歩留率（点歩留率）を定義する。
- 6) 偏差値と点歩留率の間には逆S字型の負の相関が見られ、その相関はロジスティック曲線（蓑谷，1985：186）

$$y = f / \{1 + A \exp(Bx)\} \quad (2)$$
 y ；点歩留率， x ；偏差値
 で表される（モデル曲線）。
- 7) (2)式中のパラメーターAとBの値は正規合格者と補欠繰上合格者で違いはないが、 f の値は正規合格者では $f=1$ とし、補欠繰上合格者では $f=f_0$ とする。ただし $f_0 \leq 1$ である。したがって、 f のみ異なる2本のモデル曲線の関係は図1のようになる。
- 8) 年度が異なっても、同じ偏差値なら同じ点歩留率が期待できる。

4 方法

4.1 過年度入試結果の回帰分析

以下の一連の作業はExcel（Microsoft社の表計算ソフト）の一つのシート上で行った。4.1.2と4.1.3の作業はExcel中の「ソルバー」というツールを用いて行った（川瀬・内藤，2010：170）。

4.1.1 データの収集

分析対象学科・年度の受験生の総合点の平均点（Mean）と標準偏差（SD），最高点から補欠繰上合格最低点までの各点における合格者数と入学者数を収集した。

4.1.2 回帰曲線のパラメーター（AとB）の決定

まず収集したデータから，(1)式により最高点から補欠繰上合格最低点までの総得点を偏差値に換算した（薬科学科は5.1参照）。

モデル曲線を表す(2)式中のパラメーターは， f は $f=1$ に固定して，AおよびBは正規合格者

のみのデータを用いて，(3)式

$$S = \sum_i w_i (y_{oi} - y_{ci})^2 \quad (3)$$

で計算される残差平方和(S)が最小になるようにAおよびBの値を定める最小二乗法により決定した（川瀬・内藤，2010：174）。ただし， w_i はある得点における合格者数， y_{oi} は点歩留率の実測値， y_{ci} は点歩留率の計算値である。得られたモデル曲線の実測値（ y_{oi} ）への適合度は，寄与率（ R^2 ）を文献（高橋，2005）により計算して評価した。

4.1.3 回帰曲線のパラメーター（f）の決定

パラメーター f は，正規合格者に関しては $f=1$ としたが(3の7)および4.1.2参照)，補欠繰上合格者に関しては，補欠繰上合格者のみのデータを用いて，（パラメーターAおよびBは4.1.2で得られた値に固定して）4.1.2と同様の最小二乗法により決定した。

4.2 新年度入試の入学者数予測

以下の一連の作業はExcelの一つのシート上で行った。

4.2.1 データの収集

対象学科の受験生の総合点の平均点（Mean）と標準偏差（SD），最高点から予想補欠最低点より十分低い点数まで（全収集点）の各点における人数を収集した。

4.2.2 合格者数の決定

まず収集したデータから，(1)式により全収集点の総得点を偏差値に換算した（薬科学科は5.1参照）。

次に，パラメーター f は $f=1$ とし，全収集点にわたり，4.1で得られた複数の回帰曲線を用いて予想入学者数を計算した。

ここで予想入学者数とは，最高点から任意

の得点までの累計の予想入学者数(T)のことであり、(4)式により求めた。

$$T = \sum_i n_i \times y_i \quad (4)$$

ただし、 n_i はある得点を得た受験者数、 y_i はその得点における点歩留率回帰値である。

計算結果(予想入学者数のリスト)をもとに、安全のため予想入学者数が入学してほしい人数(目標入学者数)より若干少なめになるあたりに、合格最低点(および正規合格者数)を決定した。

4.2.3 補欠数の決定

パラメーター f は、合格最低点以上は $f=1$ とし、それ未満は $f=f_0$ とし、全収集点にわたり、4.1 で得られた複数の回帰曲線を用いて、4.2.2 と同様に予想入学者数を計算した。その計算結果をもとに、安全のため予想入学者数が目標入学者数より十分多めになるあたりに、補欠最低点(および補欠者数)を決定した。補欠最低点は低くする(補欠者数が多くなる)ほど大学にとっては一見都合がいいが、受験生の立場に立てば好ましいことではなく、長い目で見れば大学にとっても好ましいことではない。

4.2.2 と 4.2.3 でのパラメーター f の切り替えは、Excel 中で IF 文を使えば簡単にできる(そのアルゴリズムは「もし総得点 \geq 合格最低点なら、 $f=1$ とし、そうでなければ $f=f_0$ とせよ」となる)。合格最低点を 4.2.2 では全収集点のいずれよりも低く設定し、4.2.3 では 4.2.2 で決まった値に設定して計算結果をパソコンから印刷すれば、入試関係者の議論の際の資料が得られる。

4.2.4 補欠からの繰上合格の決定

補欠からの繰上合格の決定は、正規合格者の入学手続状況を 4.2.3 の計算結果と見比べながら行うとよい。

5 結果と考察

5.1 受験生の質・モデル曲線・偏差値換算

受験生全体の質に関しては、評定平均値および浪人生の占める割合が参考になる(一般に浪人生は評定点の割に学力が高い)。

補欠繰上合格者の点歩留率は正規合格者の回帰曲線から予想される値より低く出る傾向にあったが、得点領域の狭さなどのため、信頼できるパラメーター A および B の値を得るのは困難であった。そこで、3 の 7) に示したように、補欠繰上合格者のモデル曲線は正規合格者のモデル曲線の上限值を押し下げた形として扱った。

薬科学科の受験者数は薬学科の受験者数の約 1/4 と少ないので、年度間の入試問題の難易度は薬学科の平均点によりよく反映されていると判断した。したがって、薬科学科の偏差値換算も薬学科の平均点(Mean)および標準偏差(SD)を用いて行った。薬科学科の結果(表 2 および図 3)はそのようにして得られたものである。

5.2 過年度入試の回帰分析結果

4.1 で述べた方法により 2008 年度および 2009 年度入試の回帰分析を行った。得られたパラメーターの値を表 2 に示す。概して B の値は安定している。それに対し、 A と f_0 の値はばらつきが大きい傾向にあり、特に薬科学科でその傾向が大きいのはデータ数の少なさによると思われる。パラメーター A および B の値の違いがどの程度回帰曲線に影響するかを知るため、図 2 および図 3 に正規合格者の偏差値域(薬学科は約 60~80; 薬科学科は約 62~75)よりも広い偏差値域にわたる回帰曲線を示す。これらの図は全偏差値域で $f=1$ として作成した。

寄与率 (R^2) は、いずれの場合もきわめて低い値を示した。ロジスティック回帰分析における R^2 の値は、通常回帰分析で得られる R^2 の値よりも低くなりがちではある (高橋, 2005) が、データ数が小さいため実測の点歩留率のバラツキが相当大きくなり R^2 の値が低くなったものとする。

表 2 回帰分析の結果 (1)寄与率; (2)3 の 7) 参照)

(2)式のパラメーター ($f=1$)					
年度	A	B	(R^2) ¹⁾	f_0 ²⁾	
薬学科	2008	3.78×10^{-4}	0.132	(0.0432)	0.993
	2009	4.18×10^{-4}	0.129	(0.0346)	0.769
	2010	8.81×10^{-4}	0.118	(0.0370)	0.942
薬科学科	2008	5.27×10^{-5}	0.160	(0.0545)	0.611
	2009	2.95×10^{-5}	0.177	(0.0435)	0.969
	2010	3.90×10^{-5}	0.167	(0.0548)	0.667

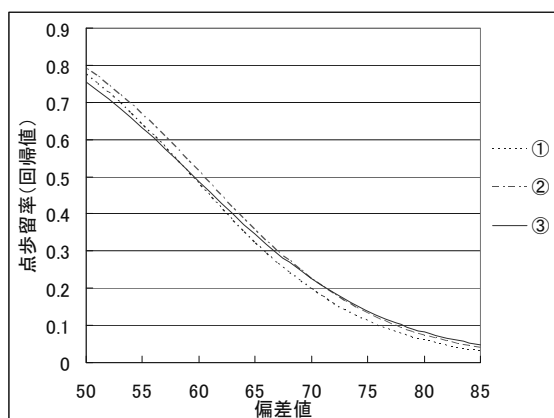


図 2 薬学科点歩留率回帰曲線

①2008 年度, ②2009 年度, ③2010 年度

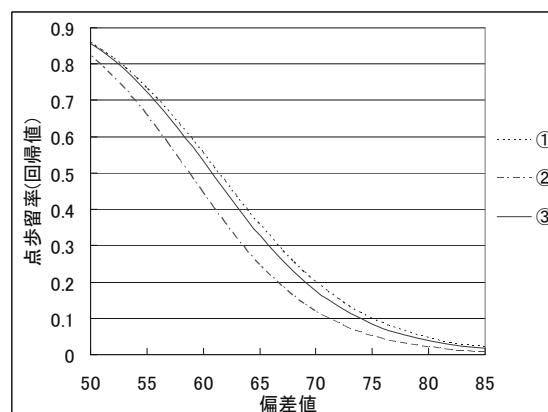


図 3 薬科学科点歩留率回帰曲線

①2008 年度, ②2009 年度, ③2010 年度

2008 年度と 2009 年度の回帰曲線を比較すると、薬学科 (図 2) ではかなり一致しているが、薬科学科 (図 3) では一致はよくない。回帰曲線の信頼性は薬学科より薬科学科のほうが低いとみるべきだろう。

薬科学科は薬学科と比べてパラメーター A が小さく B が大きいという特徴があるが、このことは図 3 (薬科学科) では図 2 (薬科学科) と比べて高偏差値域で点歩留率が低いという特徴として現れている。

5.3 新年度入試への適用結果

4.2 で述べた方法により、5.2 で得られたパラメーター値を用いて 2010 年度入試の入学者数予測を行った。一般入試 (B 方式) からの実際の入学者数は、2008 年度データからの予想入学者数と 2009 年度データからの予想入学者数の平均値と比較すると、両学科とも約 1 名のずれでありよく一致した (ずれの率としては薬科学科の方が数倍大きい)。

2010 年度の入学者が決定した後、2010 年度入試についても回帰分析を行った (結果は表 2, 図 2, 図 3 中に加えた)。薬学科のパラメーター A の値が大きくなっているのが目立つが、そのほかは 5.2 での議論の範囲内とみなせる。回帰曲線は薬学科の場合 (図 2), 2010

年度は2008年度および2009年度と偏差値域約60~80あたりではほぼ重なっている。薬科学科の場合(図3)、2010年度は2008年度と2009年度の間ながら2008年度寄りに存在している。

5.4 歩留率の意味

入学者数予測のよりどころとして、歩留率の年次推移のみを用いている大学もまだあるかもしれないので、その意味を、図1を用いて以下に論じる。

ある入試があったとしよう(受験者数(少)の曲線で説明)。合格者数を多くすれば低い偏差値(図中のa)まで合格するので歩留率は高くなる。でも、同じ入試で合格者数を少なくすれば高い偏差値(図中のb)までしか合格できないので歩留率は低くなるだろう。

一方、異なる年度間では受験者数が明らかに異なる場合はよくあることである(2つの曲線で説明)。このような場合、合格者数を同じにすると、受験者数の少ない年度(合格最低点は図中a)より受験者数の多い年度(合格最低点は図中c)の方が高い偏差値までしか合格できないので歩留率は低くなるだろう。

このように歩留率は合格者数および受験者数に左右される性質を有する。「実質倍率(=受験者数/合格者数)が上がると歩留率は下がる」という表現も、これにより説明がつく。歩留率の明らかな変化は、たいていこれにより説明がつくことを20年余りにわたる本学薬学部の入試データで確認してきた。したがって、実質倍率が毎年同じとみなせる場合は同じ歩留率を期待してよいだろうが、一般にそうみなせない場合の方が多いため、歩留率の年次推移のみから入学者数を予測するのは理にかなったやりかたではない。このようなやり方では、正規合格者数を必要以上に少なくし、補欠を必要以上に多く発表せざるをえなくなる。これでは受験生に敬遠

されるであろうし、大学も合格発表後に入学者数確保に気をもむことになる。

6 結語

入試データから比較的簡単に得られる点歩留率を用いる本回帰分析法を用いると、本学薬学部の薬学科のように合格者の偏差値域での回帰曲線の重なりが大きい場合(5.3参照)は、比較的よい精度で入学者数予測できると思われる。薬科学科のようにデータ数が少ないため、そのような回帰曲線の重なりがよくない場合には予測の精度は落ちるだろうが、回帰分析のデータが蓄積されてくれば入学者数予測への信頼性は増してくると思われる。

謝辞

本回帰分析法の価値をご理解いただき、快く入試データの閲覧を御許可いただいた本学薬学部の増野匡彦学部長に心から感謝いたします。

参考文献

- 朝日新聞社, 2008, 「大学 よめない入学者数」, 『朝日新聞』(2008年5月8日).
- 福田宏, 2003, 「経営情報学部合格者数と入学者数の関係」, 『経営と情報』, 16(1): 137-139.
- 福田宏, 2005, 「経営情報学部入学者選抜の回帰分析」, 『経営と情報』, 19(1): 39-43.
- 橋本貴充, 2007, 「偏差値を利用した簡易得点調整法——年度間比較を例として」, 『大学入試研究ジャーナル』, 17, 185-190.
- 川瀬雅也・内藤浩忠, 2010, 『化学のための数学入門』, 化学同人.
- 小林みどり・高野加代子, 2004, 「入学者数予測と合格者数決定について」, 『経営と情

報』, 16(2) : 57-60.

蓑谷千風彦, 1985, 『回帰分析のはなし』, 東京
図書.

日本私立学校振興・共済事業団 私学経営情報セ
ンター, 2010, 『平成 22 (2010) 年度 私立大
学・短期大学等入学志願動向』, 日本私立学
校振興・共済事業団.

高橋信, 2005, 『マンガでわかる統計学 回帰分
析編』, 東京図書.