

第2章 レポートの書き方

第1節 レポートの構成

レポートの形式は、対象としている分野によって多少異なったものとなるが、一般的には、最低限以下の点に留意して書き進めていくべきである。

1.1 想定される読者

レポートといっても多様である。しかしまず始めに、対象とする読者が誰なのか、はっきり意識しておく必要がある。学術誌のようにその分野の専門家なのか、より広い分野の研究者なのか、あるいは新聞などのように不特定多数なのか、学生実験のレポートのように担当教員となるのか、といった具合である。それによって、使用できる語彙群は異なったものとなるであろうし、必要に応じて脚注などの説明が必要となるかもしれない。

1.2 全体的な構成

レポートの基本的な形式を、以下のように構成する。

I. 緒論 II. 方法 III. 結果 IV. 考察 V. まとめ（必要に応じて）

「緒論」（序、はじめに、Introduction）では、当該研究を行うにあたっての背景——学術論文ならば、問題提起、これまでの研究経緯、必要性、目的など——を述べる。学生実験のレポートであれば、その実験で検討される課題の理論的背景、あるいは実験で修得されるべき技術的課題などを述べることになる。

「方法」では、取り扱われる試料（あるいは採取場所などの背景情報）についての説明、具体的な実験方法、個々の手法の理論的な説明、などを要領よくまとめる。

「結果」は、「方法」の項のどの部分に対応するか、わかりやすい形でまとめる。

「考察」では、個々の結果を素材として、緒論で述べた目的、課題などに応える形で、論考を進める。その際、結果として示された事実、そこから自分が論理的に推論した事柄、あるいは他の先行研究などによって支持（あるいは反対）される部分などの違いが、読者によって判断可能なように述べる。

結果あるいは考察において論点がいくつかに細分されたレポートでは、必要に応じて「まとめ」（おわりに、Conclusion）の項を設けてもよい。

1.3 論理構成

レポートにおいては、緒論における問題提起を考察でどのように論考していくか、あるいは個々の結果からどのように考察を進めていくか、などといった場面における論理構成が重要である。書き始める前に、まず全体の論理の流れをチャート図などで自分なりに整理しておくことが望ましい。

個々の文章は、論理的かつ要領よく書き進める。一般的に、語句の修飾関係が複雑な、長い文章は勧められない。ある文章が実験結果あるいはこれまで得られている知見を直接述べるものなのか、あるいは事実に基づいた推論であるのか、きちんと区別して書く。

実験した結果得られた事実を整理し、これら結果に基づく解釈、推論を論理的に展開していくために、図表の活用、統計的手法の適用を積極的に考えるべきである。これらの詳細については本章第2節、第3節に述べる。

第2節 データ処理

2.1 データの性質

生物実験で得られる多くのデータには、不注意な実験操作による誤差や測定装置自体の性能に由来する誤差に加えて、実験者が制御することができない様々な生物個体間および実験環境間の差異に由来する誤差が含まれる。精密な実験、熟練した研究者であれば、実験操作自体の誤差は小さくすることができるが、野生植物個体間の遺伝的な差異や実験動物個体間の発育の違いなど、研究対象の本性に由来する「ばらつき」によって生じる誤差は避けることができない。したがって、統計的データ処理は実験データから誤差（ノイズ）と「真の差異」（シグナル）を見分けるために必要不可欠である。

2.2 母集団と標本

標本とは実験や観察に用いた試料全体を指す。母集団とは実験や観察の対象となる生物種、系統、細胞全体を指す。もし、10人のヒトの赤血球数を調査してヒトの赤血球数を知りたいとすれば、選んだ10人が標本であり母集団は全人類である。同じ調査を21歳の日本の男性10人を調査対象にして行うなら、その10人が標本であり母集団は日本に住む21歳の男性に限定される。また、植物ホルモンがトマトの生育に与える影響を実験する場合は、同じ処理を無限回繰り返した時の結果全体が母集団であり、1回の実験結果が標本である。研究者は調査した10人についてのみ知りたいのではなく、また1回の実験結果だけが必要なわけではない。本当に知りたいのは母集団の赤血球数であり、植物ホルモンをトマトに処理した時に予想される影響である。

2.3 変異

標本中の個々の試料がもつ他と区別できる特性を変異と呼ぶ。また、変異のもつ情報の性質から、変異は以下のように分けることができる。

区分		情報の性質	例
定量変異	連続変異	量的な比較ができる	体重や体長
	不連続変異	量的な比較ができる	斑点や歯の数
定性変異	順序	量的な比較はできないが、順序が示せる	食べる順番や飛翔の方向
	判別（類別）	量的な比較も、順序も示せないが、異同がわかる	遺伝子の有無や反応の有無

実験者が発色や反応の強度を、5：非常に強い、4：強い、3：やや強い、2：弱い、1：非常に弱い、0：全く無い、のように数値化する場合がある。一見、定量変異のように錯覚するが、この場合の変異は量的な比較が出来るわけではないので順序である。定性変異には、後に述べる t 検定や、 F 検定などは利用できないので注意が必要である。

2.4 正確さと精度

正確さと精度は、通常同じ意味で使われるが、統計処理の中では区別される。精度（precision）とは、繰り返し測定した時に再現性のある測定値が得られる程度をいう。正確さ（accuracy）とは、測定値と真の値との近さをいう。補正されていないが非常に鋭敏な秤は、精度は高いが正確さは低いことになる。また、ある計測値はたまたま真の値に近いが、再度同

じものを計測した場合に同じ値が得られない秤は、精度が低いことになる。

2.5 測定単位と有効数字

連続変異の場合には、試料の測定値の最小単位は測定器の最小目盛に左右され、真の値は測定できないことがある。最小単位が 1mm の定規で得られた測定値 12.3cm は真の値が 12.25cm から 12.34999...cm の範囲にあることを示している。この場合の有効数字は 3 である。一般的に測定の最小単位は、最大値と最小値の間にある最小単位の数が 30 から 300 の間になるように設定する。例えば、最大値が 8mm、最小値が 4mm の場合には最小単位を 1/10mm に設定できるように測定器を選定すべきである（最小単位の数は 40 個）。逆に、最大値が 183.4cm で最小値が 36.8cm の場合には、最小単位は 1cm で十分である（最小単位の数は 146 個）。

2.6 頻度分布

得られたデータに関しては、必ず頻度分布を作成することを心掛ける。分布の左右非対称性（歪み）には、環境ストレスによる致死、天敵による捕食など生物的要因が関与する場合がある。また、分布に 2 つのピークが認められる場合には、標本の中に異質なグループ（オスとメスなど）が含まれる可能性がある。頻度分布の作成は以下の要領で行う。

- 1) 最大値と最小値をみつけ、その範囲を計算する。上の例では、4mm (=8.0-4.0), 146cm (=183-37) である。
- 2) 階級間隔を決める。範囲を 10~20 の適当な数で割る。例では、4mm/15=0.266...mm, 146cm/15=9.733...cm である。階級間隔は測定最小値単位の整数倍とするので、測定値の数が 100 より多ければ切り捨てた値を、100 よりも少なければ切り上げた値を階級間隔とする。上の例において、前者のデータ数が 54 個、後者が 237 個であれば、階級間隔は、それぞれ、0.3mm および 9cm にする。
- 3) 階級境界値を決める。最初の階級が最小値を含み測定値が 2 つの階級にまたがらないように最小単位以下の値を用いて決める。上の例では、それぞれ最初の階級境界値は 3.95-4.25mm, 36.5-45.5cm である。また、階級値を階級境界値の平均値とする。例では、それぞれ 4.1mm および 41cm である。
- 4) 度数分布表を作成する。各階級に含まれる測定値の数を度数と呼ぶ。各度数を全数に対する比率で示した場合には相対度数と呼ぶ。
- 5) 度数分布表に基づいて、頻度分布図を作成する。連続量の場合には、折れ線でも棒でも構わない。不連続量の場合には、折れ線を使うべきではなく棒と棒との間もあけておく。

2.7 平均値・中央値・最頻値

標本の測定値を代表する値として、平均値、中央値、最頻値がある。平均値は、測定値の合計/測定値の数で与えられる。平均値は測定値の分散の値を最小にする値であり、統計的データ処理に最も良く使われる。中央値は測定値を大きさ順に並べたときに真中にくる値である。つまり、標本の中にはその値より大きい試料と小さい試料が同数だけ含まれる。平均値とは異なり、一部の非常に大きくかけ離れた値に左右される程度が少ないので、分布が極端にどちらかに偏るときには標本の代表値として平均値よりも相応しい。50%発芽までの日数や半致死量などは中央値である。最頻値は観察された測定値の中で最も頻度が高い測定値である。データの分布が左右対称で単一のピークをもつ場合には、3つの値の位置は等しくなる。真の値と推定値とを区別するため、母集団の平均値を μ 、標本の平均値を \bar{x} で示す。

2.8 分散・レンジ

標本の測定値の「ばらつき」を示す尺度として、レンジと分散がある。レンジ（範囲）は最大値と最小値の差である。分散（平均平方）は平均値の周りの測定値の分布を示す。測定数が n のとき偏差平方和 S 、分散 V および標準偏差 s は以下の式で与えられる。

$$S = \sum_{i=1}^n (x_i - \bar{x})^2 \quad V = S/n - 1 \quad s = \sqrt{V}$$

このとき、分散の自由度は $f=n-1$ である。母集団の分散 σ^2 の推定値は標本の分散 V で与えられる。なお、母集団の分散および標準偏差は σ^2 および σ で、標本の分散および標準偏差は V および s で示す。

2.9 推定と検定

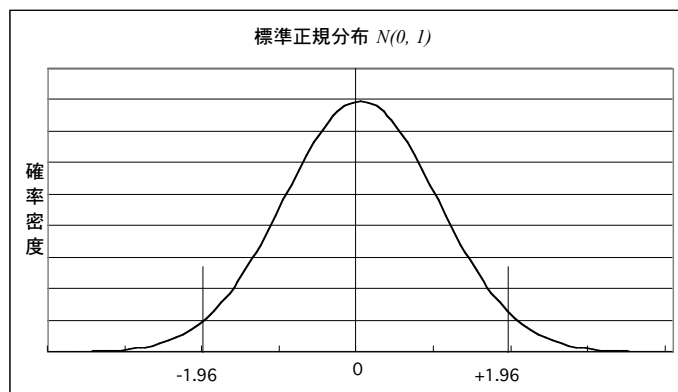
標本の値から母集団の値を導くことを推定という。標本の値と母集団との異同あるいは2標本の値の異同を調べることを検定という。統計的検定は以下の要領で行う。①比べたい2標本が同じ母集団に由来すると仮定する（帰無仮説）。②同じ母集団から2標本を取り出したときの、標本値（平均値や分散の大きさ）の差の頻度分布を考える。差が小さい場合の頻度は高く、差が大きくなる場合の頻度は小さくなるはずである。③2つの標本値から両者の間の差を求めて、この差よりも大きな差が観察される頻度を頻度分布から求める。④もし、求めた頻度が5%（もしくは1%）以下の場合には、同じ母集団に由来する標本ではないと結論する（帰無仮説の棄却）。この結論が誤っていて実際は同じ母集団に由来する可能性は5%（1%）である。結論が誤っている可能性が5%（1%）であるため、「危険率もしくは有意水準5%（1%）で差がある」とする。つまり、統計的検定とは100%「同じ」あるいは「異なる」と判定することではなく、条件付きで（危険率 α を提示して）結論することである。もし、求めた差より大きな差が観察される確率が5%以上ある場合には、「同じ母集団から由来している可能性が高い」と判断する（帰無仮説は棄却されない）。しかし、異なる母集団に由来する標本であるにも関わらず同じ母集団から由来する標本であると結論する場合の危険率は計算できない。2つの母集団の平均値や分散が予め判明していれば危険率の計算は可能であるが、それを知ることができないのは自明である。危険率が未判明のときに「同じである」と結論する統計的検定方法はない。したがって、この場合は「同じである」ではなく、「有意差がない」あるいは「差があるとはいえない」などの消極的表現を用いる。

2.10 正規分布

正規分布の確率分布密度は次式で与えられる。

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$$

式から明らかなように、分散 σ と平均値 μ が与えられれば分布は1つに決まる。任意の正規分布の変数値 x を $z=(x-\mu)/\sigma$ を用いて z に変換すれば、 z は平均値0、分散1の標準正規分布をする。図は平均値0、分散1の標準正規分布の確率密度分布を示したものである。



標準正規分布では- 1.96 から+1.96 の間に 95%の測定値が分布する。同じ母集団から標本を取ってそれぞれの平均値を求めると、平均値の分布が正規分布に近似できることが多いので、統計的検定に際してしばしば利用される。母集団の分布が厳密に正規分布をしていなくても検定は可能であるが、正規分布から外れるにしたがって検定精度は低くなることに留意する必要がある。

2.11 分散の推定と検定

母集団の分散の推定値は標本（測定数 n ）の偏差平方和を $n-1$ で割った V で与えられ、これを分散の点推定値という。母集団が正規分布であり、その分散が σ^2 である場合には $\chi^2 = S/\sigma^2$ (S は標本の偏差平方和) が自由度 $f=n-1$ の χ^2 分布をすることを利用して、標本の偏差平方和 S から母分散の信頼区間を推定することができる。すなわち、 S/σ^2 の値が χ_a^2 より小さい確率を 2.5%、 χ_b^2 より小さい確率を 97.5% とすれば、 S/σ^2 の値の 95% は 1) の範囲に分布する。 χ_a^2 および χ_b^2 の値は自由度 $f=n-1$ の χ^2 分布表から読みとればよい。1) 式を変形すれば、母集団の分散 σ^2 は 95% の確率で 2) の範囲にあると推定することができる。これを分散の区間推定という。

$$\chi_a^2(n-1, \alpha = 0.025) < \frac{S}{\sigma^2} < \chi_b^2(n-1, \alpha = 0.975) \quad \cdot \cdot \cdot 1) \quad \frac{S}{\chi_a^2} > \sigma^2 > \frac{S}{\chi_b^2} \quad \cdot \cdot \cdot 2)$$

次に、2 標本の分散の異同を検定する場合を説明する。それぞれ測定数 n_a 個と n_b 個の 2 標本を同じ母集団から抽出して算出した分散をそれぞれ V_a および V_b とする。このとき、 $F=V_a/V_b$ (ただし、 $V_a > V_b$ の時) の F 値は自由度 $f_1=n_a-1$ および $f_2=n_b-1$ の F 分布をする。両者が等しければ、 F 値は 1 であり、両者の差異が大きくなるにしたがって F 値も大きくなる。算出した F 値が $F(f_1, f_2; \text{危険率 } 5\%)$ よりも大きければ、「危険率 5% で両者は同じ母集団に由来する分散でない」、すなわち「分散の大きさが異なる」とする。 F 値は F 分布表から読みとればよい。このとき、分子は必ず大きい方の分散の値であり、 f_1 はその自由度である。

2.12 平均値の推定と検定

母集団（平均値 μ 、分散 σ^2 ）から得られた標本（測定数 n ）の平均値は、母集団の平均値の点推定値であり、平均値 μ 、分散 σ^2/n の正規分布をする。この分布を標準正規分布に変換すると、その値の 95% は 3) の範囲に分布する。したがって、母平均 μ は確率 95% で 4) の区間に存在すると推定できる。

$$-1.96 < \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} < +1.96 \quad \cdot \cdot \cdot 3) \quad \bar{x} - 1.96 \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + 1.96 \frac{\sigma}{\sqrt{n}} \quad \cdot \cdot \cdot 4)$$

通常は母集団の分散 σ^2 は未知であるため、標本の標準偏差 s を用いて 5) の t 値を算出する。 z 値と t 値とは似ているが、 t 値は標準正規分布ではなく自由度 $f=n-1$ (測定数 n) の t 分布をする。したがって、母平均 μ が確率 95% で存在する範囲は、1.96 の代わりに $t(f, \alpha)$ を用いて示した 6) の範囲であると推定できる。 $t(f, \alpha)$ 値は t 表から読みとればよい。 α は危険率である。

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} \quad \cdot \cdot \cdot 5) \quad \bar{x} - t(f, \alpha = 0.05) < \mu < \bar{x} + t(f, \alpha = 0.05) \quad \cdot \cdot \cdot 6)$$

2 つの標本の平均値 x_a および x_b の異同を検定するためには、同じ母集団（平均値 μ 、分散 σ^2 ）から抽出したときに、両者の差異 ($x_a - x_b$) の分布が 7) の正規分布をすることを利用して、8) のように t 値を算出する。7) に標本から得られた平均値と $\mu_a - \mu_b$ の期待値 0 を代入して計算

し、この値が $t(n_a+n_b-2, \alpha=0.05)$ と比較して大きければ、「危険率 5% で x_a-x_b の値は 0 ではない」、すなわち「危険率 5% で両者は異なる」とする。

平均値 0, 分散 $\frac{\sigma^2}{\left(\frac{1}{n_a} + \frac{1}{n_b}\right)}$ (・・・7)

$$t = \frac{(\bar{x}_a - \bar{x}_b) - (\mu_a - \mu_b)}{s_d \sqrt{\frac{1}{n_a} + \frac{1}{n_b}}} \quad \text{自由度 } f=n_a+n_b-2 \quad \text{・・・8)} \quad \text{ただし, } s_d = \sqrt{\frac{S_a + S_b}{(n_a - 1) + (n_b - 1)}}$$

この検定では両標本の分散 V が等しいことを前提とするので、両標本の分散に差異が無いことを F 検定で予め確認しておく。分散が等しくない場合には補正法を用いるが検定精度は低くなる。

第3節 図表の書き方

3.1 図表とは

実験で得られた結果を簡潔に視覚化して示したり，考察に説得力を与えたりする上で，図表の利用は欠かせない。しかし，適切に利用しなければ論旨を混乱させたり，誤解を生むので注意する。図は，数値データを棒・折れ線・点などに変換して適切な座標軸上に配置するもの（数値図）と，ひとまとまりの文言を論旨がわかりやすいように配置するもの（フローチャートなど）があるが，ここでは実験結果の提示に利用されることの多い数値図について説明する。表は，いくつかの行と列を持つ表に数値データや文字を整理して並べて提示する。図表に共通の注意事項について記したあと，それぞれについて特筆すべき事項について説明する。

3.2 図表と本文の関係

自分が言いたいことを説明するのに必要な図表のみを，必要な情報のみを使用して提示する。したがって，図表について本文中で必ず言及する。ただし，例えば「実験結果は表1の通りであった。」だけでは，読者が得られる情報はない。表1のどのデータに注目して欲しいか，表1が示していることは何か，本文中で説明する。

3.3 図表の体裁

図表は，それ自体で提示されている内容が理解できるように整える必要がある。そのために，

- I. 番号と題を付す。題には，図表の内容が理解しやすいものをつける。
- II. 提示したデータには単位をつける。
- III. 必要であれば，欄外に図表の注釈をつける。

以下3.4では，仮想のデータとして1950年に森林を伐採し耕地とした場所での植生および土壌中の炭素賦存量の気候帯別の経時的変化を例に用いて説明する。

3.4 図の種類と書き方

図には，棒グラフ，折れ線グラフ，散布図，円グラフなどがある。図の題は，図の下に付

【棒グラフ】

比較するのに適したグラフである（図1）。横軸には比較対象となる項目を並べ，縦軸には比較したい数量をとる。また，図1にあるように比較対象項目の内訳を分けて（植物と土壌）棒の積み重ねで表現することもできる。

【折れ線グラフ】

時間的な推移や傾向を表現するのに適したグラフである（図2）。横軸には時間を，縦軸には数量をとる。図2にあるように複数の系列を同時に提示するときには，各系列に異なるマークを用いるとともに，そのマークがどの系列に当たるかを凡例として示す。

【散布図】

異なる数値データの間を関係性を提示するのに適したグラフである（図3）。横軸・縦軸ともに数量をとる。図3は，土壌炭素賦存量の減少率が年平均土壌温度と関係があることを示している。

【円グラフ】

全体を100として各項目の比率を示したいときに使用する（図4）。図4では，気候帯別に植生の部位別炭素賦存量を示し，円の中心には植生全体の炭素賦存量を記している。このグラフは，構成割合について説明したいとき（この場合は，気候帯ごとに部位別の炭素賦存量

の割合が異なることを示したいとき)に使用すべきであり,分母が異なること(植生帯ごとに植物中の全炭素賦存量が異なること)を強調したいときには,積み重ねの棒グラフを使う.

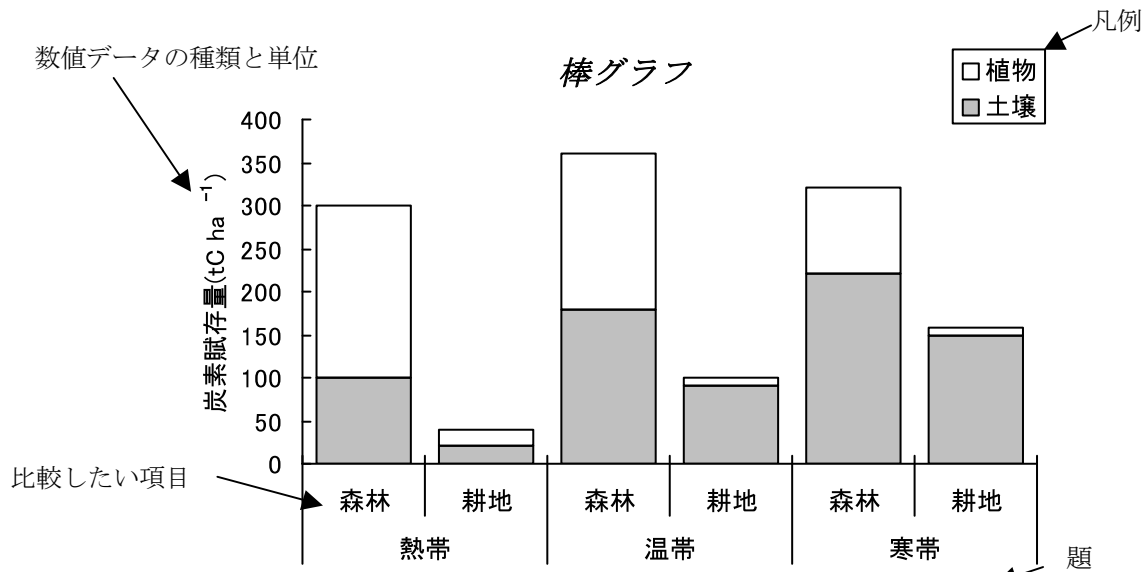


図1 気候帯・土地利用別の植物と土壌中の炭素賦存量
 森林は19 50年の森林伐採時の炭素賦存量、
 耕地は20 00年の炭素賦存量である。

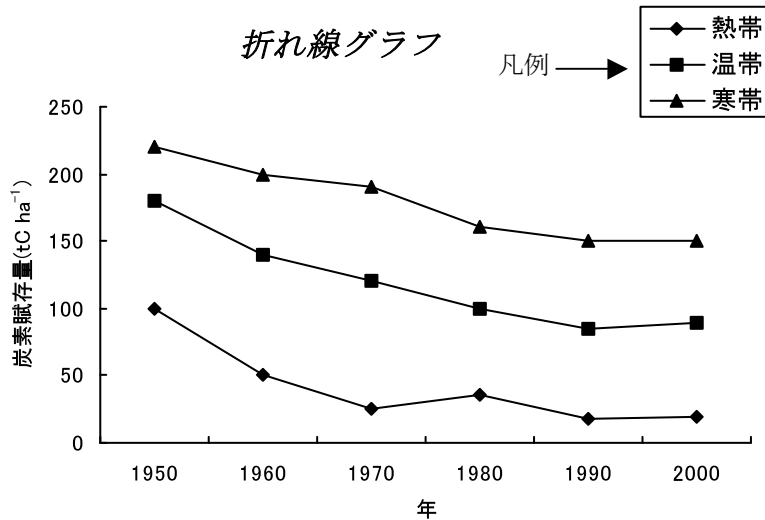


図2 気候帯別の森林伐採後の土壌炭素賦存量の変化

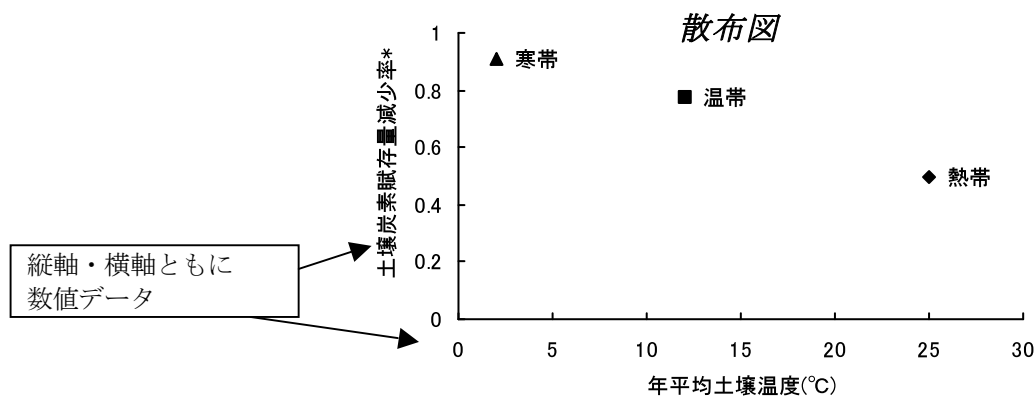


図3 年平均土壌温度と森林伐採直後10年の土壌炭素賦存量減少率の関係

*: 土壌炭素賦存量減少率
 = 森林伐採直前(1950年)の土壌炭素賦存量
 ÷ 伐採後10年(1960年)の土壌炭素賦存量

数値データの詳しい説明が必要な場合は注釈をつける

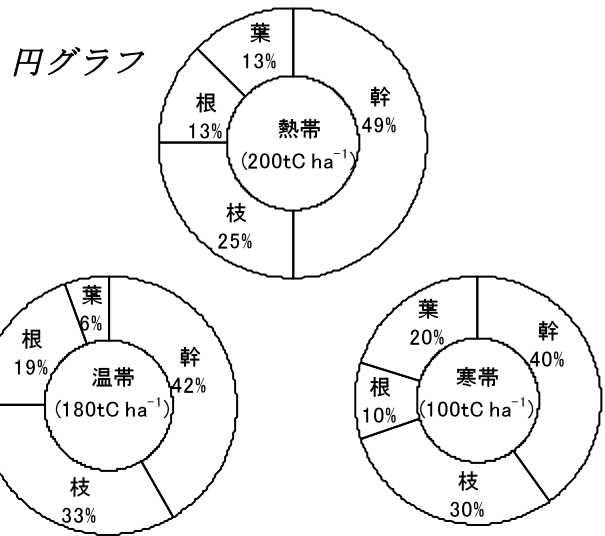


図4 気候帯別の植物中の炭素賦存量の部位別割合
()内の数字は、植物全体の炭素賦存量を示す。

3.5 表の体裁

表は、文章で羅列して説明すると理解しにくいものや、単位が異なるため図では表現しにくい場合に使用する。表の題は、表の上に付す。表内に数値を記すときは、該当する比較項目に単位を付す。

比較項目で表の行を構成する

比較対象で表の列を構成する

表1 調査地点の概要

	国	年平均土壌温度(°C)	年平均降水量(mm)	自然植生のタイプ
熱帯	インドネシア	25	4000	熱帯雨林
温帯	日本	12	2500	照葉樹林
寒帯	カナダ	2	500	タイガ

行

列