

経済統計b:第九回

担当教員 黒田敏史

今週の内容

- アンケート調査
 - アンケート作成・回収方法
 - アンケートの集計・分析

今週の内容

- アンケート調査
 - アンケート作成・回収方法
 - アンケートの集計・分析

アンケート調査の仕方

- アンケート調査の意義
 - 社会科学の対称となる様々な現象を考察するにあたって、通説・印象と実際の姿が一致している保証はない
 - 社会科学の対称となる現象では、実験によってデータが得られるとは限らない
 - ミクロ経済学の対称となる、個人の行動に関するデータは実験で得やすい
 - 他方、企業行動や、観察対象となると行動が変わるもの等は実験でデータを得ることが難しい
 - アンケート調査によってデータを収集、分析する事で、科学的な研究を行う事ができる

アンケート調査の仕方

- アンケート調査の流れ
 - 目的の設定
 - 対象の設定
 - 調査手法の設定
 - 調査実行・回収
 - 分析

アンケート調査の仕方

- 目的の設定

- アンケートを行う前に、何のためにアンケートを行うのかという目的を明確にする必要がある
 - 最終目的を明確にしなければ、どの様なアンケートを、何時までに行うか、等を決めることができない
 - 作成途中で迷った際には、最終目的を実現するためにもっとも適した手法を考えれば良い(逆向き帰納法)
 - アンケートよりも好ましい手法があるかもしれないし、アンケートが不要かもしれない
- 例:「経済学の常識と大学生の常識を比較したい」

アンケート調査の仕方

- 対象の設定

- アンケートをとる対象を、調査したい内容に応じて設定する必要がある

- 真の姿を知るには母集団の全標本から回答を得るのが理想だが、費用・期間等から実行不可能なことが多い

- 調査対象の検討項目

- 構成単位：主体は個人・世帯・企業・社員等に分類できる

- 必要条件：調査を行う相手が満たしているべき条件はあるか（年齢・性別・経験、世帯構成、業種・企業規模、勤務先・役職）

- 時間的・空間的条件：居住地（都道府県・市町村・人口規模別・通勤圏、居住年数、設立年数、勤務年数）

- 先の例の場合：大学生（経済学部生とその他）

アンケート調査の仕方

- 対象の設定

- サンプルングの実施

- 母集団から一部を抽出する調査を標本調査と呼ぶ

- 母集団のリストがある場合

- 無作為抽出: リストからランダムに抽出
 - 系統抽出: リストから抽出する最初の標本だけを無作為に決め、2番目以下の標本を一定の間隔で決める
 - 層化抽出: 母集団を網羅・背反なグループに分け、グループ毎に無作為抽出する(標本が偏りにくくなる)

- 母集団のリストがない場合

- エリアサンプリング: 地図からつまみ上げる、等
 - タイムサンプリング: 通行人の5人に1人を拾い上げる、等

アンケート調査の仕方

- 調査手法の設定

- アンケートには様々な手法がある

- 留め置き法・郵送法・集合法・面接・電話・インターネット
…等

- アンケートの分類

- 記入者による分類: 自記式・他記式
 - メディアによる分類: 面接・電話・インターネット・テレビ等
 - タイミングによる分類: 留め置き・即時・パネル
 - 場所による分類: 来場者・訪問

アンケート調査の仕方

- 調査手法の設定

- アンケート手法の比較項目

- サンプルング関連

- 調査対象リストが必要か
 - サンプルングバイアスは防げるか
 - 対象を捉えやすいか

- 調査員関連

- 調査員は適切な技術を有しているか
 - 調査員の教育、管理、監督が必要か

- コスト

- 費用(金銭やその他)
 - 調査にかかる期間はどの程度か

- 回答精度関連

- 調査への理解・協力が得やすいか
 - 回答が確実に得られるか
 - 正直な回答が得られるか
 - 複雑な質問について説明出来るか
 - 対象者が理解力に劣る場合でもフォローできるか
 - 記入ミスが防止できるか
 - 商品見本や道具などを使った質問が可能か

アンケート調査の仕方

独断と偏見による調査手法比較

	リストの必要性	統計的な代表性	調査地域	回答精度			期間	費用	回収率(括弧内は世論調査のもの)	その他
				量	正直さ	理解度				
訪問面接調査	あり	だいたいできる	拡大で費用が大きく増大	▲	▲	○	中	高	高(80%)	調査員を訓練する必要がある
訪問留め置き調査	あり	だいたいできる	拡大で費用が大きく増大	○	▲	○	長	高	高(70-80%)	調査員を訓練する必要がある
郵送調査	あり	だいたいできる	範囲と費用は無関係	○	○	▲	長	中	中(50-60%)	督促を行う必要がある
電話調査	あり	できることもある	範囲と費用はあまり関係ない	×	○	○	短	中	中(70%)	調査員を訓練する必要がある
インターネット調査	あり	できることもある	範囲と費用は無関係	○	▲	▲	短	低	低(-)	
来場者調査	なし	来場者に限定される	範囲と費用はあまり関係ない	▲	▲	▲	短	低	低(-)	会場があるか?
ホームユーステスト	あり	モニタに限定される	範囲と費用はあまり関係ない	▲	○	○	長	高	高(-)	一般にない経験をしてもらえる

アンケート調査の仕方

- 調査手法の設定

- アンケートの作成

- アンケートの構成を決定する

- 調査の説明: アンケートの目的・実施主体・調査対象・結果の利用方法等を説明する

- 質問項目: 目的となる質問を行う

- » 回答に必要な所要時間等を勘案して、設問数や選択肢数を決定する

- フェイスシート: 回答者の属性を記入してもらう

- » 年齢・性別・職業・居住地、等々

- » 不必要に詳細な属性を求めると、回答に協力してもらえなくなる場合があるため、最小限に留める事が良いとされている

アンケート調査の仕方

- 調査手法の設定

- アンケートの作成

- 設問と回答のタイプ

- 自由回答型質問: 回答者に自由に回答をしてもらう

- » 「税の負担を実感するのはどの様な時ですか 答_____」

- » 「消費税率は何%以下が好ましいですか 答____%」

- プリコード型質問: 予め用意した回答リストから回答を選択してもらう

- » 例: 「選択肢 1・重い 2・普通 3・軽い」

- 自由回答付きプリコード: プリコードと自由回答両方を用いる

- » 例: 「選択肢 1・良い 2・普通 3・悪い 4・その他(自由回答)」

- » 例: 「選択肢 1・好き 2・嫌い 3・どちらでもない その理由(自由回答)」

アンケート調査の仕方

- 調査手法の設定

- アンケートの作成

- 設問の作成における注意点

- 質問文を簡潔にする: 回答者の負担を軽減する
 - 質問文を明確にする: 回答者が設問の意味を理解出来るようにする
 - 誘導的な質問文にしない: 関連価値・ラベリング・現状肯定
 - 一つの質問で複数のことを聞かない
 - » 悪い例: 「1・美容と健康によい 2・甘くて美味しい」
 - 質問相手を明確にする
 - » 例: 「あなたが」、「家庭で」、「身の回りで」、「日本で」
 - 手法に応じた質問文にする
 - » 口頭と文章、媒体の違いで適切な表現・情報量は変わってくる

アンケート調査の仕方

- 調査手法の設定
 - アンケートの作成
 - 自由回答の特徴
 - 利点
 - » 質問者の予期していない回答が得られる
 - 欠点
 - » 数値を直接記入させるときの回答精度は概して低い
 - » 集計を行う際の負担が大きい

アンケート調査の仕方

- 調査手法の設定

- アンケートの作成

- プリコードのタイプ

- 単一回答

- » 例:「次の中からもっともよく利用しているサービスをお答えください。

- 1・メール 2・電話 3・Web 4・音楽 5・動画」

- » 例:「食べ残しの持ち帰りに賛成ですか

- 1・賛成 2・反対」

- 複数回答

- » 例:「次の中から利用中のサービスを全てお答えください

- 1・メール 2・電話 3・Web 4・音楽 5・動画」

- » 例:「〇〇に当てはまるイメージを全てお答えください

- 1・信頼できる 2・地味 3・有料 4・教育 5・総合」

アンケート調査の仕方

- 調査手法の設定
 - アンケートの作成
 - 選択肢の尺度
 - 名義尺度: 性別・職業・地域などに番号を割り当てる
 - 順序尺度: 比較可能なランクに番号を割り当てる
 - » 1・とても良い 2・良い 3・普通 4・悪い
 - 間隔尺度: 等間隔なランクに番号を割り当てる
 - » 1・とても良い 2・良い 3・悪い 4・とても悪い
 - 比例尺度: 基数性のある項目に番号を割り当てる
 - » 1・0～9冊 2・10～19冊 3・20～29冊
 - 名義尺度・順序尺度は、その設問の中での相対的な関係の分析を行ったり、グルーピングを行うのに用いる(割合の計算、属性によってグループを絞り込む等)
 - 例: 所属学部は何処ですか 1・経済 2・経営 3・コミュ 4・法
 - 間隔尺度や比例尺度であれば、平均値等を計算し、相関をとる等の他の設問との関連を分析することができる
 - 間隔尺度の点数をどの様に割り当てるかに明確な指針はない
 - 例: 取得単位数はいくつですか

アンケート調査の仕方

- 調査手法の設定

- アンケートの作成

- コーディングの注意点

- 用意した選択肢が回答の集合を網羅している事が好ましい

- » 悪い例:「1・とても賛成 2・賛成 3・どちらかと言えば賛成」

- » 悪い例:「ご利用の携帯電話はどれですか

- 1・ドコモ 2・au 3・ソフトバンク

- » 悪い例:「1・美容と健康によい 2・甘くて美味しい」

- 用意した選択肢が排他である事が好ましい

- » 悪い例:「1・1,000円以下 2・1,000-2,000円 3・2,000-3,000円 4・それ以上」

- » 悪い例:「1・東京 2・京都 3・関西 4・名古屋 5・九州」

- 選択肢の選択基準が明確であることが好ましい

- » 悪い例:「重視する要素はどれですか 1・ブランド 2・社名 3・イメージ 4・技術」

- » 悪い例:「1・北海道 2・東北 3・関東 4・中部 5・北陸 6・近畿 7・中国 8・四国 9・九州 10・沖縄」

アンケート調査の仕方

- 調査実行・回収
 - 設定したサンプリング法に応じて実際のアンケートを収集する
 - サンプリングが良くできていても、対象者が不在であったり、回答拒否が行われる事によって調査結果に偏りが生じる事がある

今週の内容

- アンケート調査
 - アンケート作成・回収方法
 - アンケートの集計・分析

アンケートの集計・分析

- 調査実行・回収
 - 回収データのチェック
 - 回収したデータに不正や虚偽申請がないかをチェックする
 - 調査を2重に行う、電話で確認をする、等で調査を行う
 - 回答のパターンからでたらめな回答を取り除く
 - » 同じ選択肢ばかり選ぶ、規則的な回答をしている、等
 - » 事前に設問を工夫し、複数の設問を組み合わせ、真の報告を行っていない回答者をあぶり出す事も有効
 - » これらのパターンによるフィルタリングを行ったとしても、一定割合のでたらめな回答や無矛盾な虚偽報告を取り除くことはできない
 - » 事後的なフィルタリングは、分析者の意図によるバイアスを生むため、調査設計段階で可能な限りの工夫を行い、事後的なフィルタリングは最小に留めるべき

アンケートの集計・分析

- 調査実行・回収
 - 集計作業
 - 目標数のアンケートを収集した後、アンケートの集計作業を行う
 - Excel等の表計算ソフトを用いると集計が行いやすい
 - 集計の例

Q21. 移動体通信端末を通じて現在利用しているサービスをすべてお答えください。

番号	電話(音声のみ)	テレビ電話	電子メール	Webブラウジング(iモード、Ezweb、インターネット等)	SNS/ブログ	静止画像送受信	動画送受信	音楽ダウンロード・再生	動画ダウンロード・再生	ゲームダウンロード・再生	電子書籍ダウンロード・閲覧	テレビ放送受信(ワンセグ)	ラジオ放送受信	GPS/ナビゲーション機能	国際ローミング機能	男女(男=0)	
1	1	0	1	1	1	1	1	1	0	0	0	1	0	1	0	2	
2	1	0	1	1	1	1	1	1	1	1	1	0	1	0	1	0	2
3	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1
4																	2
5	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
6	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
7	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	2
8	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
9	1	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0	2
10	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
11	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
12	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
13	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2
14	1	0	1	1	0	0	0	1	1	0	0	1	0	1	0	0	1

アンケートの集計・分析

- アンケートの分析

- 単純集計

- 設問毎の回答数をカウントする

- 単一回答の場合: 円グラフによる回答比率等を用いる
 - 複数回答の場合: 棒グラフによる回答数の比較等を用いる
 - 自由回答の場合、回答のリスト化、分類等を行う

- 記述統計による分析

- 平均値や分散を計算する

$$\text{平均} = \frac{\sum_N x_i}{N} \quad \text{不偏分散} = \frac{\sum_N (x_i - \bar{x})^2}{N-1}$$

- 属性毎のグルーピングを行い、グループ毎の傾向や、グループ間の平均や散らばりの差を見る

- » 例: 男女の違い・学部の違い・年齢の違い

アンケートの集計・分析

- アンケートの分析

- より進んだ統計分析

- 母平均の区間推定

- 標本から得られた平均から、母平均の信頼区間を求めることができる
 - 母平均の信頼区間: ある母平均から、観察された標本が得られる確率を元に、母平均がありそうな範囲を絞り込む
 - 正規分布で95%の信頼区間の場合、(正確には1.96ではなく、t分布表を用いるべき)

$$\text{標本平均} - 1.96 \sqrt{\frac{\text{不偏分散}}{\text{標本数}}} \leq \text{母平均} \leq \text{標本平均} + 1.96 \sqrt{\frac{\text{不偏分散}}{\text{標本数}}}$$

- » 信頼区間よりも外に母平均がある場合はまず無いだろうと考え、母平均が信頼区間より外の場合が無いと判断することができる(「棄却する」という)

アンケートの集計・分析

- アンケートの分析
 - 区間推定の例
 - 先の携帯電話利用データから、母平均・母分散・95%信頼区間を推定

	電話(音声のみ)	テレビ電話	電子メール	Webブラウジング(iモード、Ezweb、インターネット等)	SNS/ブログ	静止画像送受信	動画送受信	音楽ダウンロード・再生
平均	1	0.076923	0.923077	0.461538	0.153846	0.153846	0.153846	0.230769
分散	0	0.076923	0.076923	0.269231	0.141026	0.141026	0.141026	0.192308
上限	1	0.227692	1.073846	0.743602	0.357989	0.357989	0.357989	0.469156
下限	1	-0.07385	0.772308	0.179475	-0.0503	-0.0503	-0.0503	-0.00762

	動画ダウンロード・再生	ゲームダウンロード・再生	電子書籍ダウンロード・閲覧	テレビ放送受信(ワンセグ)	ラジオ放送受信	GPS/ナビゲーション機能	国際ローミング機能	その他	男女(男=0)
平均	0.153846	0.076923	0	0.230769	0	0.307692	0	0	0.384615
分散	0.141026	0.076923	0	0.192308	0	0.230769	0	0	0.25641
上限	0.357989	0.227692	0	0.469156	0	0.568832	0	0	0.659881
下限	-0.0503	-0.07385	0	-0.00762	0	0.046552	0	0	0.10935

アンケートの集計・分析

- アンケートの分析

- より進んだ統計分析

- 平均値の差の検定

- あるグループの平均値と、他のグループの平均値が、取得した標本においてたまたま異なっていたのか、それとも母平均が異なっているのかを検証する
 - T検定 (tは検定統計量、t分布に従う。s_jはグループjの分散)
 - 自由度 ($N_A + N_B - 2$) にもよるが、概ね1.96以上の場合に母集団に差があると言える (エクセルの分析ツールを用いれば容易に計算ができる)

$$t = \frac{\text{グループAの平均} - \text{グループBの平均}}{\sqrt{\frac{\text{グループAの分散}}{\text{グループAの標本数}} + \frac{\text{グループBの分散}}{\text{グループBの標本数}}}}$$

アンケートの集計・分析

• アンケートの分析

• 平均値の差の検定の例

– 男女間で携帯電話のサービス利用に差があるかを検定する

平均値の差の検定	電話(音声のみ)	テレビ電話	電子メール	Webブラウジング	SNS/ブログ	静止画像送受信	動画送受信
男平均	1	0.125	0.875	0.25	0	0	0
女平均	1	0	1	0.8	0.4	0.4	0.4
男分散	0	0.125	0.125	0.214285714	0	0	0
女分散	0	0	0	0.2	0.3	0.3	0.3
t統計量	#DIV/0!	1	-1	-2.128241472	-1.632993162	-1.632993162	-1.632993162
統計的な男女の差	不明	棄却	棄却	有り	棄却	棄却	棄却

平均値の差の検定	動画ダウンロード・再生	ゲームダウンロード・再生	電子書籍ダウンロード・閲覧	テレビ放送受信(ワンセグ)	ラジオ放送受信	GPS/ナビゲーション機能	国際ローミング機能
男平均	0.125	0	0	0.125	0	0.125	0
女平均	0.2	0.2	0	0.4	0	0.6	0
男分散	0.125	0	0	0.125	0	0.125	0
女分散	0.2	0.2	0	0.3	0	0.3	0
t統計量	-0.318	-1	#DIV/0!	-1	#DIV/0!	-1.727272727	#DIV/0!
統計的な男女の差	棄却	棄却	不明	棄却	不明	棄却	不明

– この場合、Web利用のみ女性がよく利用していると言える

アンケートの集計・分析

・ アンケートの分析

－ 全アンケートの有効回答2011サンプルによる分析

	電話(音声のみ)	テレビ電話	電子メール	Webブラウジング	SNS/ブログ	静止画像送受信	動画送受信
男平均	0.948840927	0.07833733	0.877697842	0.493205436	0.057553957	0.214228617	0.072741807
女平均	0.971052632	0.100396301	0.918421053	0.527704485	0.112285337	0.30474934	0.101717305
男分散	0.048580655	0.072258353	0.107430216	0.250153797	0.054284892	0.168469384	0.067504396
女分散	0.028146453	0.090039055	0.075022537	0.248170305	0.099364852	0.211353304	0.091089234
t統計量	-2.550068281	-1.661658365	-2.997358635	-1.503511425	-4.147549584	-4.45550556	-2.197798065
統計的な男女の差	有意	棄却	有意	棄却	有意	有意	有意
	動画ダウンロード・再生	ゲームダウンロード・再生	電子書籍ダウンロード・閲覧	テレビ放送受信(ワンセグ)	ラジオ放送受信	GPS/ナビゲーション機能	国際ローミング機能
男平均	0.076738609	0.095923261	0.028776978	0.284572342	0.029576339	0.136690647	0.049560352
女平均	0.085865258	0.124174373	0.034346103	0.275725594	0.034300792	0.117569353	0.039630119
男分散	0.070906475	0.086791367	0.027971223	0.203753797	0.02872454	0.118100719	0.047141807
女分散	0.078256442	0.108407577	0.03307439	0.199237552	0.033077505	0.103418788	0.037953076
t統計量	-0.722318007	-1.940206694	-0.686146177	0.429116495	-0.579417971	1.259500688	1.060849922
統計的な男女の差	棄却	棄却	棄却	棄却	棄却	棄却	棄却

－ この場合、電話・メールSNS/ブログ・静止画像送受信・動画送受信、ゲームに関して女性がよく利用していると言える

≫ 標本数が増えたため、分散の推定値が小さくなり、より信頼性の高い分析が可能となったことによる

次週の内容

- テキスト12章「対外バランス統計」