

2022학년도 2학기

경영 통계학

담당교수: 백수정



학습 목표

1. 통계학을 정의하고 통계학 학습의 필요성을 이해함
2. 경영통계학의 활용 범위를 살펴보고 실무에서의 당면과제를 이해함
3. 일반적인 통계 활용의 함정을 살펴봄

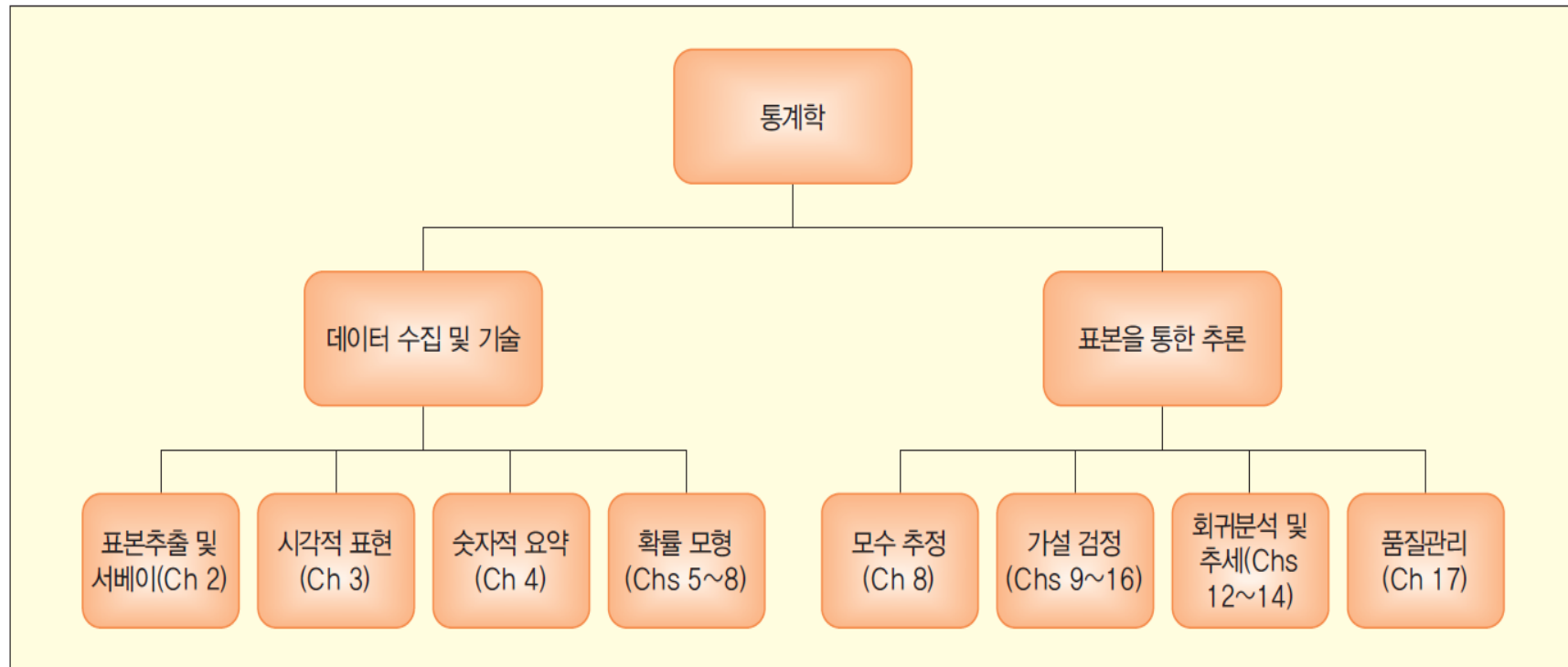
통계학이란 무엇인가?

통계학은 현실의 측정치들을 요약하고, 적절한 결론을 도출하고, 미래를 예측하는 방법들을 모아 놓은 것으로 생각할 수 있다. 통계학은 원자료를 조직화해서 정보를 제시하고 의미를 이끌어낼 수 있게 도와준다.

단수? 복수?

통계학(statistics)은 데이터를 모으고, 정리하고, 분석하고, 해석하여 그 결과를 제시하는 학문.

통계량(statistic)은 표본 데이터세트를 요약하는 데 사용되는 단일 측정수단으로서 숫자로 표현됨; 예를 들어, 어떤 대학 재학생의 평균 키.



비즈니스 애널리틱스란?

애널리틱스(analytics)는 통계, 수학, 전산 도구를 사용하여 데이터에서 정보를 추출하는 광범위한 분야를 가리킨다. 애널리틱스의 도구는 크게 기술, 예측, 규범의 세 가지 범주 이 용어는 우리가 답하려고 하는 다음의 질문에서 파생.

어떤 일이 있었는가? 기업은 기술 애널리틱스(descriptive analytics) 도구를 사용함으로써 역사적 데이터를 분석하고 경향 및 패턴을 찾아내는 데 도움을 얻는다.

다음에 어떤 일이 일어날까? 기업은 예측 널리틱스(predictive analytics) 도구를 사용함으로써 미래 사건의 가능성을 예측하고 소비자 행동 예측에 도움을 얻는다.

목표달성을 위해 어떤 조치를 취해야 할까? 기업은 규범 애널리틱스(prescriptive analytics)도구를 사용함

왜 통계학을 공부하는가?

우리는 통계학을 배움으로써 다른 사람들이 제시하는 데이터 분석을 더 잘 이해할 수 있다.

충분한 통계지식을 갖고 있어야만 다른 사람들이 과장된 주장으로 우리를 기만하는 것을 막을 수 있으며, 또한 우리 스스로 자신의 전문지식의 한계가 어디까지인지 알 수 있다.

회사 차원에서 보면 통계학 지식을 보유한 회사는 내부 및 외부 데이터를 활용하지 못하는 조직에 비해 비교우위를 가질 수 있다.

개인적으로도 우리는 기초 통계학을 마스터함으로써 관리자로서 비교우위를 갖고 더 빠른 승진의 기회를 잡을 수 있으며, 또는 더 나은 일자리를 찾을 수 있다.

왜 통계학을 공부하는가?

의사소통

통계학 전문 용어들은 언어 장벽을 극복하고 국경을 초월한 문제해결에 도움을 준다.

컴퓨터 숙련도

데이터를 분석하기 위해 스프레드시트를 다루고, 보고서를 작성하고, 프레젠테이션을 함으로써 우리는 자신이 갖고 있는 기술들을 종합하는 능력을 높인다. **정보관리**

통계학은 많은 데이터를 요약하고 거기에 담겨 있는 관련성을 밝히는 데 유용하다.

전문지식

첨단기술의 성장산업이 최고 직종이 되곤 한다. 통계 소프트웨어의 활용은 첨단 기술의 전문지식을 높여준다.

왜 통계학을 공부하는가?

품질개선

거대 제조업체인 보잉이나 도요타와 같은 경우 공식적인 상시품질개선 시스템이 가동되고 있다. 기업들은 통계학을 이용하여 납품업체 관리, 내부운영 모니터링, 문제점 발견 등에서 도움을 얻을 수 있다. 품질개선이 통계학만으로 되는 것은 아니지만 이러한 분야에서 통계학이 어떤 역할을 하고 있는지 이해할 수 있을 정도의 통계학은 알아두어야 할 것이다.

비즈니스와 통계학

감사

어떤 기업이 일부 청구서의 대금지불에 오류가 있다는 것을 알고 있지만 이 문제가 어느 정도 심각한지에 대해서는 알지 못한다. 모든 청구서를 검토할 자원이 부족하기 때문에 표본을 추출하여 오류가 발생하는 비율을 추정할 수 있다.

마케팅

많은 회사에서 고객관계관리(CRM)를 사용하여 다양한 소스의 고객 데이터를 분석한다. 상관관계나 데이터마이닝과 같은 통계 및 분석 도구를 사용하여 다양한 고객 그룹의 특정 니즈를 파악함으로써 제품 및 서비스를 보다 효과적으로 마케팅 할 수 있다.

의료

100명의 방문 환자에 대해 육체적 정신적 상태에 관한 42개 문항의 설문조사를 실시.

품질관리

3중 검사프로그램을 도입하고, 품질이 낮은 제품을 생산한 근로자에게는 벌칙을 부과.

비즈니스와 통계학

구매

어느 식품업체에 납품된 500개의 플라스틱 용기 중 3개에서 결함이 발견되었다. 제조업체가 밝힌 경험적 결함률은 0.005이다. 이번 결과로 볼 때 결함률이 상승했는가, 아니면 운이 나쁜 것인가?

의약품

신약이 위약에 비해 효과가 더 좋은지, 아니면 우연히 그런 것인지 결정한다.

운영관리

재고 관리를 위한 소비자 수요 예측.

제품 워런티

어떤 자동차 제조업체는 새로운 하이브리드 엔진에 대한 워런티 요구에 평균적으로 어느 정도의 비용이 소요되는지 알고 싶다.

통계학적 도전

이상적인 데이터 분석가(비즈니스 분야에서 통계학을 다루는 전문가)는 다음의 특성을 지녀야 한다.

- 높은 기술성
- 합리적이고 관계성 높은 의사소통
- 주도적
- 전체를 조망하는 시각
- 유연함
- 마감기한 준수
- 중요 문제에 집중
- 자신의 한계를 잘 알고 도움을 요청
- 불완전 정보를 다룰 수 있는 능력
- 전문가적 성실성

불완전 데이터와 현실적 제약

수집된 자료의 형태와 질에 있어서 제약에 당면하게 될 것이다.

가정이나 분석의 한계를 명확히 서술한다. 비정상적인 데이터 포인트를 포착하고, 결측치(missing data)를 처리하기 위해 통계적 테스트를 이용할 수도 있다.

제약의 사례

- 자동차 충돌실험에서는 사람을 쓸수 없다(위험).
- 전화면접에서 여성에게 낙태경험을 물을 수 없다(민감한 질문).
- 모든 사람을 대상으로 에이즈 검사를 할 수 없다(전세계가 실험실이 아님)

비즈니스 윤리와 통계

비즈니스와 관련된 광범위한 윤리적 책임은 다음과 같다.

- 고객을 공정하고 정직하게 대우.
- 차별을 금지하는 법률을 준수.
- 제품 및 서비스가 안전 규정을 준수하도록 보장.
- 사실적이고 유익한 방식으로 광고를 제공.
- 회사는 자신의 근로자들이 회사의 비즈니스 관행에 대해 질문을 하거나 우려를 표명하는 것을 장려.
- 개별 구성원은 경영진에게 정보를 정확하게 보고.

비즈니스 윤리와 통계

데이터 분석가를 위한 윤리규범:

- 허용된 절차를 알고 따른다.
- 데이터상의 정직성을 유지한다.
- 정확한 계산을 수행한다.
- 절차를 성실하게 보고한다.
- 비밀정보를 보호한다.
- 출처를 밝힌다.
- 재정 지원의 출처를 확인한다.

팀으로 작업하기

비즈니스 활동은 종종 팀 단위로 처리된다. 오늘날 기술 전문가들은 많은 시간을 회의하고 프로젝트 일정을 토론하며, 팀 보고서를 준비하고 방법론에 대해 논하면서 보낸다.

팀이 당면한 문제

- 팀을 구성하는 최선의 방법은 무엇인가?
- 팀이 사용하게 될 의사소통 채널은 무엇인가?
- 팀의 자원을 어떻게 잘 활용할 수 있을까?
- 팀리더의 역할은 무엇인가?
- 소위원회가 필요한가?

컨설턴트 활용하기

어떤 중요한 전문기술이 부족하거나 편향되지 않은 견해가 필요한 경우에는 어떤 프로젝트의 시작단계에서 컨설턴트를 고용할 필요가 있다.

요즘 기업들은 외부 컨설턴트가 있더라도 내부의 직원들이 모든 유형의 통계 분석을 수행할 수 있기를 기대한다.

숫자로 소통하기

숫자는 특정한 상황 속에서 소통이 될 때만 의미를 갖는다.

좋은 그래프가 전문 보고서를 훨씬 돋보이게 해주고 다른 관리자들로 하여금 필요한 정보들을 금방 이해하고 올바른 의사결정을 하는 데 도움을 줄 수 있다.

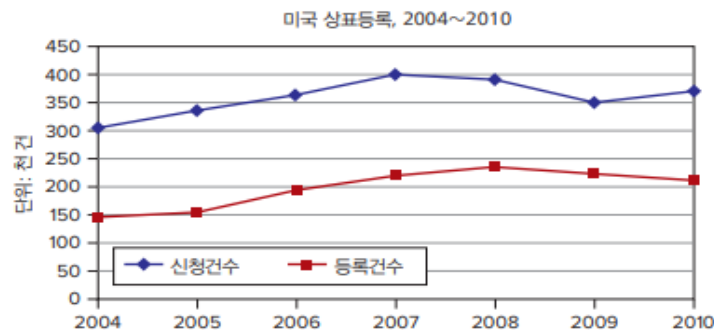
표 1.1

미국 상표등록, 2004~2010(단위: 천 건)

	2004	2005	2006	2007	2008	2009	2010
신청건수	304.5	334.7	362.3	401.0	390.8	351.9	370.2
등록건수	146.0	154.8	193.7	218.8	233.9	222.1	210.6

출처: U.S. Census Bureau, *Statistical Abstract of the United States*, 2012, p. 778.

그림 1.2

미국 상표등록
2004~2010

비판적 사고

통계학은 어떤 아이디어를 경험적 증거에 비추어 검증할 수 있기 때문에 비판적 사고(critical thinking)의 핵심 부분이다.

우리는 통계학적 도구들을 이용하여 경험적 데이터(empirical data, 관찰 및 실험을 통해 수집된 데이터)와 이론을 비교한다.

이 과정에서 빠지기 쉬운 논리적 함정이 많다.

논리적 함정

데이터를 모으는 과정과 추론하는 과정에는 많은 논리적 함정이 존재한다.

함정 1: 작은 표본에 기초하여 결론 내리기

표본이 적을 때는 일반화에 주의를 해야 한다(예를 들어 10명의 환자가 개선을 보일 때)

함정 2: 비임의(nonrandom) 표본에 의해 결론 내리기

예: 심장마비가 발생하지 않은 사람들은 조사하지 않고, 심장마비가 발생한 사람들만을 대상으로 연구하여 그것을 일반화하는 것은 조심해야 한다.

함정 3: 드문 사례로 결론 내리기

모집단 전체로 보았을 때 그다지 놀랄만한 사건이 아닌데도 강한 추론을 도출하는 경우가 있다(예: 로또 당첨).

논리적 함정

함정 4: 잘못된 서베이 방법

표본추출 방법이 잘못되었거나 질문의 표현이 잘못될 수 있다(무기명 서베이나 퀴즈시험이 더 나은 방법일 수 있음)

함정 5: 인과관계에 대한 가정

인과관계가 없는데도 있는 것처럼 결론을 내리지 않도록 주의해야 한다. 예를 들어 ‘야구장의 저주’라는 것이 있다. 즉, 이름을 가지고 있는 야구장을 홈으로 하는 팀들(예를 들어 Citi Field의 뉴욕 메츠)은 승보다는 패가 더 많다는 것이다. 그러나 어느 팀이 이기고 지는 것은 선수와 감독에 달렸다.

함정 6: 개인으로 일반화

“남자가 여자보다 키가 더 크다.” 통계적으로만 그렇다. 남자가 평균적으로 더 크지만 남자보다 더 큰 여자도 많다. 따라서 사실들을 너무 통계적 일반화(statistical generalizations)로 몰아가는 것을 피해야 한다.

논리적 함정

함정 7: 의도하지 않은 편향

분명한 사기(데이터 조작)는 아니더라도 부지불식중에 데이터를 어느 한쪽으로 몰아갈 수 있다. 예: 심장마비는 남성에게 더 쉽게 발생한다고 믿어 왔지만, 이는 여성보다 남성의 심장병 증상이 더 뚜렷하기 때문이다.

함정 8: 유의성 대 중요성

통계적으로는 의미 있는 효과라 하더라도 현실적으로는 별 의미가 없을 수 있다. 예를 들어 어떤 표본 중 봄에 태어난 사람의 키가 가을에 태어난 사람에 비해 0.6cm 더 크다. 이것은 중요한 차이일까?