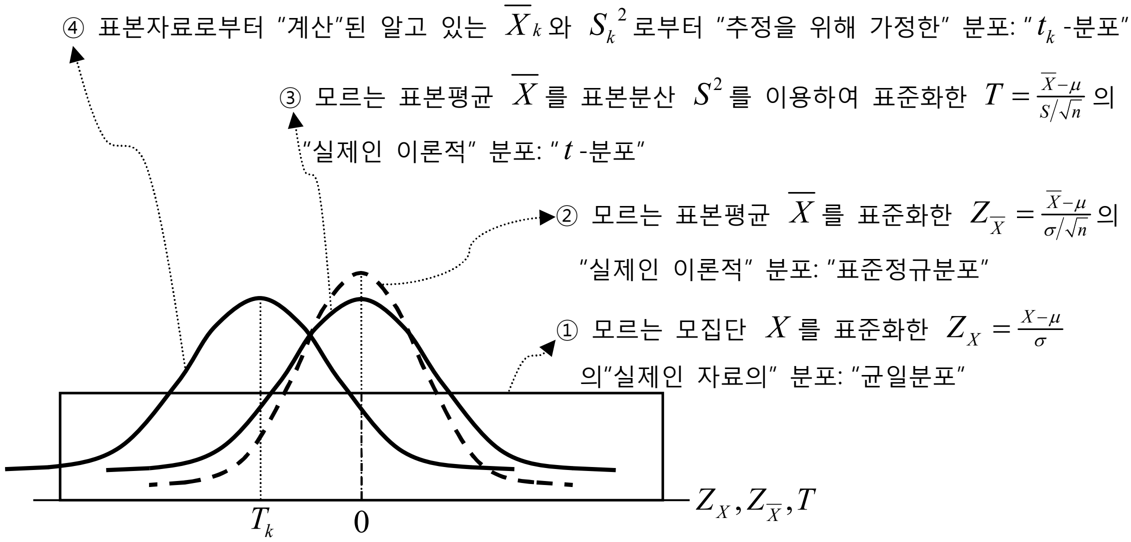
표본평균($\overline{X}$)으로부터 모평균(μ)을 신뢰수준 $1-\alpha$ 에서 "구간추정"하면 다음과 같다.

- 수식: $P(\overline{X}-t_{\alpha/2}\times\frac{S}{\sqrt{n}}\le\mu\le\overline{X}+t_{\alpha/2}\times\frac{S}{\sqrt{n}})=1-\alpha$
- 해석: $[\overline{X}-t_{\alpha/2}\times\frac{S}{\sqrt{n}},\ \overline{X}+t_{\alpha/2}\times\frac{S}{\sqrt{n}}]$가 μ를 포함하고 있을 확률은 $1-\alpha$ 이다.

그런데 위 해석을 (상식적으로는 더 익숙한 표현인) "μ가 $[\overline{X}-t_{\alpha/2}\times\frac{S}{\sqrt{n}},\ \overline{X}+t_{\alpha/2}\times\frac{S}{\sqrt{n}}]$에 포함되어 있을 확률은 $1-\alpha$ 이다"라고 해석(해석 2)하면 안된다고 한다. 왜냐하면 μ는 (확률변수가 아니라) 상수이기 때문이다.

하지만 t-분표가 대칭적이기 때문에 다음과 같은 근거로 "해석 2"로 해도 무방하다고 본다.

□ 구간추정의 의미를 이해하기 위해서는, 아래 그림에 있는 4가지 분포를 구분하여 이해해야 한다.



단, $T_k=\frac{\overline{X}_k-\mu}{S_k/\sqrt{n}}$로서 표본자료로부터 "계산"된 $\overline{X}_k$와 S_k^2가 표준화된 "값"

- 분포 ①은 "모르는" 모집단 X를 "표준화"한 $Z_X=\frac{X-\mu}{\sigma}$의 "실제인 자료의 분포"이다.
  - ▫ 표본으로부터 모집단을 추정하는 것이 목적이므로, 우리는 모집단 X의 분포를 모른다.
  - ▫ 모집단 X의 분포를 (다양할 수 있지만 여기서는) '균일분포(uniform distribution)'로 가정한다.
    - · 즉, 모집단 X의 모든 값에 대하여 그 값의 "(확률이 아니라) 도수"는 모두 동일하다.
    - · 모집단 X의 분포는 이론적(확률적) 분포가 아니라 실제 자료의 분포이다.
  - ▫ X의 "(기대값이 아닌) 평균"은 μ라고, "(확률적 분산이 아닌) 실제 분산"은 σ^2이라고 가정한다. 물론 우리는 μ와 σ^2을 "모르지만", 이것들은 "사실(true) 즉 실제"이다.
  - ▫ 평균이 μ이고 분산이 σ^2인 균일분포를 따르는 X를 표준화한 $Z_X=\frac{X-\mu}{\sigma}$도 "분포 ①"처럼 평균이 0인 균일분포가 된다.
- 분포 ②는 "모르는" 표본평균 $\overline{X}$를 "표준화"한 $Z_{\overline{X}}=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$의 "실제인 이론(확률)적 분포"이다.

▫ 비록 모집단을 균일분포로 가정했지만, 표본수 n 이 충분히 크면 중심극한정리에 의해, “이론적”으로 $Z_{\overline{X}} = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ 이다. (즉, $\overline{X}$ 를 표준화한 $Z_{\overline{X}}$ 는 표준정규분포를 따른다.)

· $Z_{\overline{X}}$ 가 확률변수이므로, 분포 ②는 이론적인 분포이다.

▫ 물론 우리는 μ 와 σ^2 을 “모르지만”, 이것들 그리고 이것들에 의해 결정되는 표준정규분포는 “사실(true) 즉 실제”이다.

– “분포 ③”은 모르는 표본평균 $\overline{X}$ 를 표본분산 S^2 을 이용하여 표준화한 $T = \frac{\overline{X}-\mu}{S/\sqrt{n}}$ 의 “실제인 이론적(확률적) 분포”이다.

▫ 표본평균 $\overline{X}$ 와 표본표준편차 S 모두 확률변수이므로, 표준화된 T 도 확률변수이다.

▫ 모집단이 균일분포이지만, 표본수 n 이 충분히 크면 “실험적”으로 $T = \frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$ 이다. ($\overline{X}$ 와 S 로 표준화한 T 는 “t-분포”를 따른다.) T 가 확률변수이므로 분포 ③은 이론적 분포이다.

▫ 물론 우리는 (다양한 값을 가질 수 있는) $\overline{X}$ 와 S 를 “모르지만”, 이것들 그리고 이것들에 의해 결정되는 t-분포는 “사실(true) 즉 실제”이다.

– “분포 ④”는 표본자료로부터 “계산”된 알고 있는 $\overline{X}_k$ 와 S_k 로부터, μ 의 신뢰구간을 “추정 혹은 해석하기 위해” (어디까지나) 편의적·기술적으로 “가정한 분포”이다.

▫ 분포 ④는 표본자료로부터 실제로 계산된 $\overline{X}_k$ 와 S_k 로부터 도출된 “$T_k = \dfrac{\overline{X}_k - \mu}{S_k/\sqrt{n}}$ 를 중심으로 한 그리고 μ 를 확률변수로 취한” t-분포와 형태가 동일한 분포로서, 분포 ③을 T_k 가 중심이 되도록 (분포의 형태는 유지한 채) 이동시킨 (어디까지나) “가상”의 분포이다.

· $T_k = \frac{\overline{X}_k-\mu}{S_k/\sqrt{n}}$ 는 (어디까지나) “분포 ③”에 해당하는 확률변수인 $T = \frac{\overline{X}-\mu}{S/\sqrt{n}}$ 가 취할 수 있는 값 중의 하나이다. 그러므로, 변수가 아닌 값인 T_k 는 사실 “분포를 가질 수 없다”.

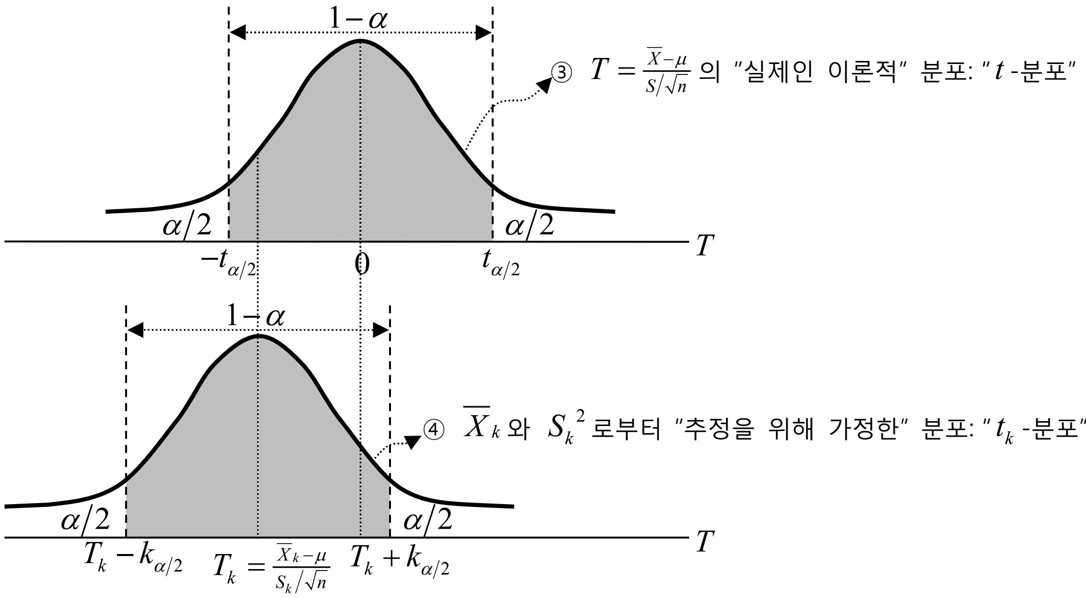
· 하지만 관점을 바꾸어 (모르는) μ 를 확률변수로 취하면 T_k 는 분포를 가지게 된다.

· 앞의 그림에서는 $T_k < 0$ 즉 $\mu > \overline{X}_k$ 으로 가정하여 분포 ④를 분포 ③의 왼쪽에 그렸지만, T_k 의 값에 따라 ($T_k = 0$ 즉 $\mu = \overline{X}_k$ 이면) 분포 ③과 같을 수도 있고 ($T_k > 0$ 즉 $\mu < \overline{X}_k$ 이면) 분포 ③의 오른쪽에 있을 수도 있다.

▫ “t_k-분포”라고 명한 이유는 중심만 다를 뿐 t-분포와 형태가 동일하기 때문이다.

· t_k-분포는 일종의 “유사(pseudo) t-분포”이며, 중심인 T_k 의 수만큼 존재한다.

□ 아래 그림은 "분포 ③"과 "분포 ④"를 동일한 T값을 기준으로 (두 분포를 겹치게 그리는 것이 맞지만) 구분하기 쉽게 각각 독립적으로 그려 놓은 것이다.



– (앞 그림의 위쪽) "분포 ③"에서 짙은 색 영역은 (원래 의미대로) 확률변수 $\overline{X}$와 S가 상수 μ를 기준으로 한 $P(-t_{\alpha/2}\le\frac{\overline{X}-\mu}{S/\sqrt{n}}\le t_{\alpha/2})=1-\alpha$ 임을 나타내고 있다.

▫ 즉, "t-분포"의 기대값(중심) 0을 기준으로 대칭인 두 값 $-t_{\alpha/2}$과 $t_{\alpha/2}$ 사이에 $T=\frac{\overline{X}-\mu}{S/\sqrt{n}}$의 값이 포함될 확률은 $1-\alpha$ 이다.

– (앞 그림의 아래쪽) "분포 ④"에서 짙은 색 영역은 (어디까지나 가상적 의미로) 확률변수 μ가 상수 $\overline{X}_k$와 S_k를 기준으로 한 $P(T_k-k_{\alpha/2}\le\frac{\overline{X}_k-\mu}{S_k/\sqrt{n}}\le T_k+k_{\alpha/2})=1-\alpha$ 임을 나타내고 있다.

▫ 즉, "t_k-분포"의 기대값(중심) $T_k(=\frac{\overline{X}_k-\mu}{S_k/\sqrt{n}})$ 를 기준으로 대칭인 두 값 $T_k-k_{\alpha/2}$ 과 $T_k+k_{\alpha/2}$ 사이에 $T_k(=\frac{\overline{X}_k-\mu}{S_k/\sqrt{n}})$의 값이 있을 확률은 $1-\alpha$ 이다.

▫ 즉, 분포 ③은 (원래대로) $\overline{X}$와 S가 확률변수이고 μ가 상수인 반면, 분포 ④는 반대로 μ가 확률변수이고 $\overline{X}_k$와 S_k가 상수인 것으로 (가상적으로) 취급하고 있다는 점에서 서로 다르다.

– 이제 이 과정의 핵심인 "분포 ③"과 "분포 ④"의 확률적 관계를 살펴보자.

▫ 분포 ③의 "t-분포"는 기대값 0을 중심으로 "대칭"인 분포이다.

▫ 분포 ④의 "t_k-분포"는 (앞에서 가상으로 정의하였듯이) T_k를 중심으로 분포 ③의 t-분포와 형태가 동일한 "유사(pseudo) t-분포"이다. 그러므로 분포 ④도 "대칭"인 분포이다.

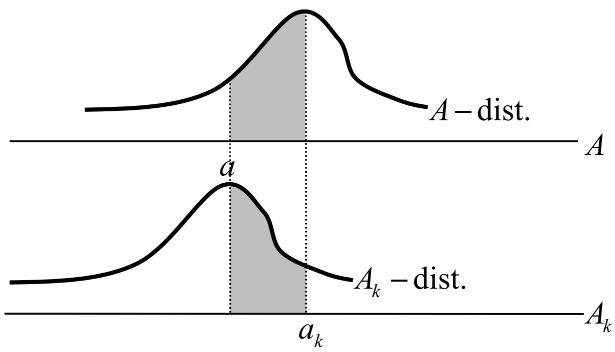
▫ 그러므로 (앞 그림의 아래쪽) t_k-분포의 짙은 색 영역이 t-분포의 기대값 0을 포함할 확률과, (앞 그림의 위쪽) t-분포의 짙은 색 영역이 t_k-분포의 기대값 $T_k=\frac{\overline{X}_k-\mu}{S_k/\sqrt{n}}$를 포함할 확률은 서로 동일하다.

· "t-분포"의 기대값 0은 단순한 숫자 0이 아니라 $E(\frac{\overline{X}-\mu}{S/\sqrt{n}})=0$ 혹은 $\mu=\overline{X}$ 임을 의미한다.

▫ 즉, $P(T_k-k_{\alpha/2}\le 0\le T_k+k_{\alpha/2}\,|\,t_k-\text{dist.})=P(-t_{\alpha/2}\le T_k\le t_{\alpha/2}\,|\,t-\text{dist.})$ 이다.

· 단, 이 결과는 분포 ③과 분포 ④가 중심을 제외하면 다른 "형태는 모두 동일"하면서, 또한 두 분포 모두 각자가 "대칭분포"이기 때문이다.

· 예를 들어, 아래 그림처럼 분포가 비대칭이면 $P(a_k\le a\,|\,A_k-\text{dist.})\ne P(a_k\le a\,|\,A-\text{dist.})$ 이다.



– 이제 위의 결과를 이용하여 "현재 우리의 목적인" μ의 구간을 도출해보자.

▫ 신뢰수준 $1-\alpha$에서 μ의 신뢰구간은 가상의 분포 ④로부터 다음 식으로 도출된다.

$P(T_k-k_{\alpha/2}\le 0\le T_k+k_{\alpha/2}\,|\,t_k-\text{dist.})$

▫ 그런데 위 식은 (앞에서 살펴보았듯이) 분포 ③에 해당하는 다음 식과 동일하다.

$P(-t_{\alpha/2}\le T_k\le t_{\alpha/2}\,|\,t-\text{dist.})$

▫ 분포 ③에 해당하는 후자의 식 즉 $P(-t_{\alpha/2}\le \frac{\overline{X}_k-\mu}{S_k/\sqrt{n}}\le t_{\alpha/2}\,|\,t-\text{dist.})$를 이용하면,

$P(\overline{X}_k-t_{\alpha/2}\times\frac{S_k}{\sqrt{n}}\le\mu\le\overline{X}_k+t_{\alpha/2}\times\frac{S_k}{\sqrt{n}})=1-\alpha$ 가 도출된다.

$$\begin{aligned}\because P(-t_{\alpha/2}&\le\frac{\overline{X}_k-\mu}{S_k/\sqrt{n}}\le t_{\alpha/2})\\ &=P(-t_{\alpha/2}\times\frac{S_k}{\sqrt{n}}\le\overline{X}_k-\mu\le t_{\alpha/2}\times\frac{S_k}{\sqrt{n}})\\ &=P(-\overline{X}_k-t_{\alpha/2}\times\frac{S_k}{\sqrt{n}}\le-\mu\le-\overline{X}_k+t_{\alpha/2}\times\frac{S_k}{\sqrt{n}})\\ &=P(\overline{X}_k-t_{\alpha/2}\times\frac{S_k}{\sqrt{n}}\le\mu\le\overline{X}_k+t_{\alpha/2}\times\frac{S_k}{\sqrt{n}})\end{aligned}$$

▫ 즉, 신뢰수준이 $1-\alpha$이면 μ의 신뢰구간은 $[\overline{X}-t_{\alpha/2}\times\frac{S}{\sqrt{n}},\overline{X}+t_{\alpha/2}\times\frac{S}{\sqrt{n}}]$가 된다.

지금까지 살펴본 바에 따르면 "μ가 구간 $[\overline{X}-t_{\alpha/2}\times\frac{S}{\sqrt{n}},\overline{X}+t_{\alpha/2}\times\frac{S}{\sqrt{n}}]$에 포함되어 있을 확률은 $1-\alpha$이다"로 해석(해석 2)하는 것이 잘못되었다는 기존 주장이 오히려 잘못된 것일 수도 있음을 의미한다.