Part IV 노출평가 연구

AI 산업보건 석사논문

9주차· 10주차· 11주차

이 Part 에서 답할 질문

- 노출평가 연구는 무엇을 재고, 무엇을 답하는가 — "평균이 기준 이하"가 왜 답이 아닌가
- 실제 측정은 어떻게 설계하는가 — SEG, 배경 농도, 비교 가능한 조건
- 비검출과 단위는 왜 분석 코드가 아니라 정의서가 정하는가

0. 문제로 시작합니다

용접 공정의 작업환경측정 결과가 왔습니다. 시료 49개, 평균 3.9 mg/m³, 노출기준 5 mg/m³. 보건관리자가 묻습니다. "평균이 기준 아래니까 괜찮은 거지요? 그런데 지난달 국소배기를 설치했는데 효과가 있었는지도 알고 싶습니다. 그리고 조립 라인은 절반이 '불검출'인데 이건 어떻게 씁니까?"

세 질문 모두 평균으로는 답할 수 없습니다. 노출은 오른쪽 꼬리가 긴 분포이고, 개선 효과는 같은 조건에서만 비교되며, 불검출은 0 이 아닙니다.

1. 무엇인가 — 1.1 정의와 목적

노출평가는 **근로자가 유해인자에 얼마나 노출되는가를 측정·추정**하는 연구입니다. 답하려는 것은 셋입니다 — 이 집단의 노출 분포는 어떤가, 노출기준을 얼마나 자주 넘는가(판정), 무엇이 노출을 바꾸는가(공정·환기·개선 효과). 역학 연구의 노출 변수(Part I 의 소음 dB)가 여기서 나옵니다.

1. 무엇인가 — 1.2 무엇을 재는가

갈래	무엇	예
작업환경측정	공기 중 농도. **개인시료**(호흡대, 8h TWA·STEL) · 지역시료	용접흄, 유기용제, 소음
생물학적 모니터링	체내 지표. 채취 시점이 결과를 정함	요중 대사산물, 혈중 납
노출 복원·모델링	잴 수 없었던 과거·미래 노출	직무노출매트릭스(JEM), 몬테카를로

1. 무엇인가 — 1.3 언제 쓰는가

법정 측정의 해석, 개선 조치의 우선순위, 역학 연구의 노출 변수 생성, 기준 개정의 근거(용량–반응). 학위논문으로는 공공 측정자료 분석, 측정 전략 비교, 노출 모델링이 한 학기에 가능합니다. 실제 시료 채취·실험실 분석은 장비가 필요해 학기 안에 어렵습니다.

1. 무엇인가 — 1.4 예

용접 공정의 SEG 별 흄 노출과 초과 확률 · 국소배기 설치 전후 비교 · 도장 작업 유기용제의 계절 추세 · 소음 노출과 청력 손실(Part I 의 노출 변수).

2. 기본 구조

요소	내용
분석 단위	**시료** 하나 = 한 행. 근로자·SEG·날짜·시료 시간·농도·단위·비검출 여부
집단	**유사노출그룹(SEG)** — 같은 공정·작업·환기 → 하나의 분포
분포	**대수정규** — GM(기하평균)·GSD(기하표준편차)로 기술
판정 지표	95% UCL(평균의 상한), **기준 초과 확률**, 95 백분위수
산출물	SEG 정의표 · 시료 요약표(n·비검출·GM·GSD·UCL·P>OEL·판정) · 분포 그림 · 비교 표(배경·전후·추세)
참조	노출기준(OEL)은 **고시 원문**에서. 단위(mg/m³ / ppm) 명시

3. 하는 법 단계별 — 3.1 측정 설계 — 사람이 현장에서 정합니다

측정 설계는 사람이 현장에서 정하는 일 — 분석은 이 여섯 결정을 바꾸지 못한다

① 무엇을 재는가
개인시료(호흡대) · 8h TWA
지역시료는 개인 노출이 아님

② 누구를 묶는가 — SEG
같은 공정·작업·환기
근거를 문서로

③ 몇 개를 재는가
SEG 당 6~10 이상
비검출 비율 예상

④ 배경 농도
작업 전·비작업 구역·사무실
작업 기여 = 작업 − 배경

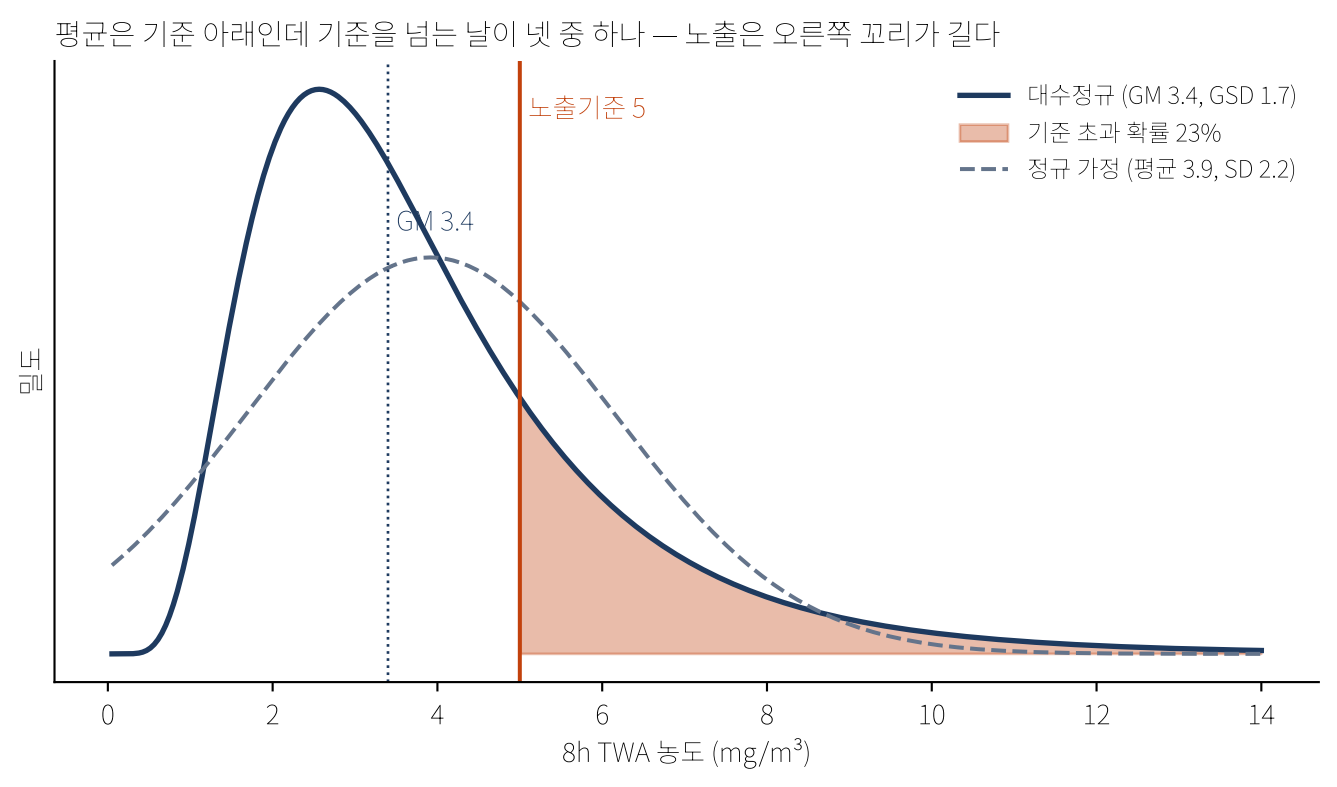
⑤ 비교 가능한 조건
개선 전후는 같은 SEG·계절·
생산량에서만 비교

⑥ 기준값과 단위
고시 원문에서 · mg/m^3/ppm
정의서가 정한다

3. 하는 법 단계별 — 3.1 측정 설계 — 사람이 현장에서 정합니다

단계	내용	틀리면
① 무엇을	개인시료 8h TWA (또는 STEL). 지역시료는 개인 노출을 말하지 못함	지역시료로 판정
② SEG	공정·작업·환기가 같은 사람을 묶는다. 근거 문서화	이질 집단을 한 분포로
③ 시료 수	SEG 당 6~10 이상. 비검출 비율 예상	UCL 이 n 을 무시하지 않는다
④ 배경 농도	작업 전·비작업 구역·사무실 시료. 작업 기여 = 작업 − 배경	배경이 높은 분진에서 작업 효과 과대
⑤ 비교 가능한 조건	개선 전후는 같은 SEG·계절·생산량에서만	생산량이 준 시기와 비교해 "개선 효과"
⑥ 채취·보정	펌프 유량 보정, 360분 이상, 채취 기록	15분 시료로 TWA

3. 하는 법 단계별 — 3.2 분포 기술 — 대수정규



3. 하는 법 단계별 — 3.2 분포 기술 — 대수정규

농도에 로그를 취하면 정규분포에 가까워집니다. 로그 평균의 지수가 **GM**, 로그 표준편차의 지수가 **GSD** 입니다. GSD 2 이면 "GM 의 2배 / 1/2배 안에 약 68%". 기준 초과 확률은 $P(X > OEL) = 1 - \Phi\left(\frac{\ln OEL - \ln GM}{\ln GSD}\right)$ 로 계산합니다. 95% UCL 은 산술평균의 상한으로, 대수정규에서는 Land 방법 또는 그 근사를 씁니다 — t 구간이 아닙니다.

3. 하는 법 단계별 — 3.2 분포 기술 — 대수정규

"평균이 기준 이하"가 아니라

"기준 초과 확률이 얼마"를 보고합니다

3. 하는 법 단계별 — 3.3 비검출(LOD 미만)

방법	언제	편향
삭제	절대 안 됨	위로
0 대치	안 됨	아래로
LOD/2, LOD/√2	비검출 < 30%	작음
좌측 중도절단 MLE · ROS	비검출 30~80%	권장
모수 보고 안 함	비검출 > 80%	"기준 대비 낮음"까지만

3. 하는 법 단계별 — 3.4 판정과 비교

- 판정 — UCL 과 초과 확률을 조치 기준(예: 초과 확률 5%)과 비교. "허용 / 조치 필요 / 부적합"
- 배경 대비 — 작업 시료와 배경 시료의 분포 차이
- 전후 비교 — GM 비와 CI. **조건이 같았는지**를 먼저 표로 (생산량·계절·환기)
- 추세 — 월별 로그 농도 회귀. 인접 달은 닮아 있어 **자기상관**을 고려

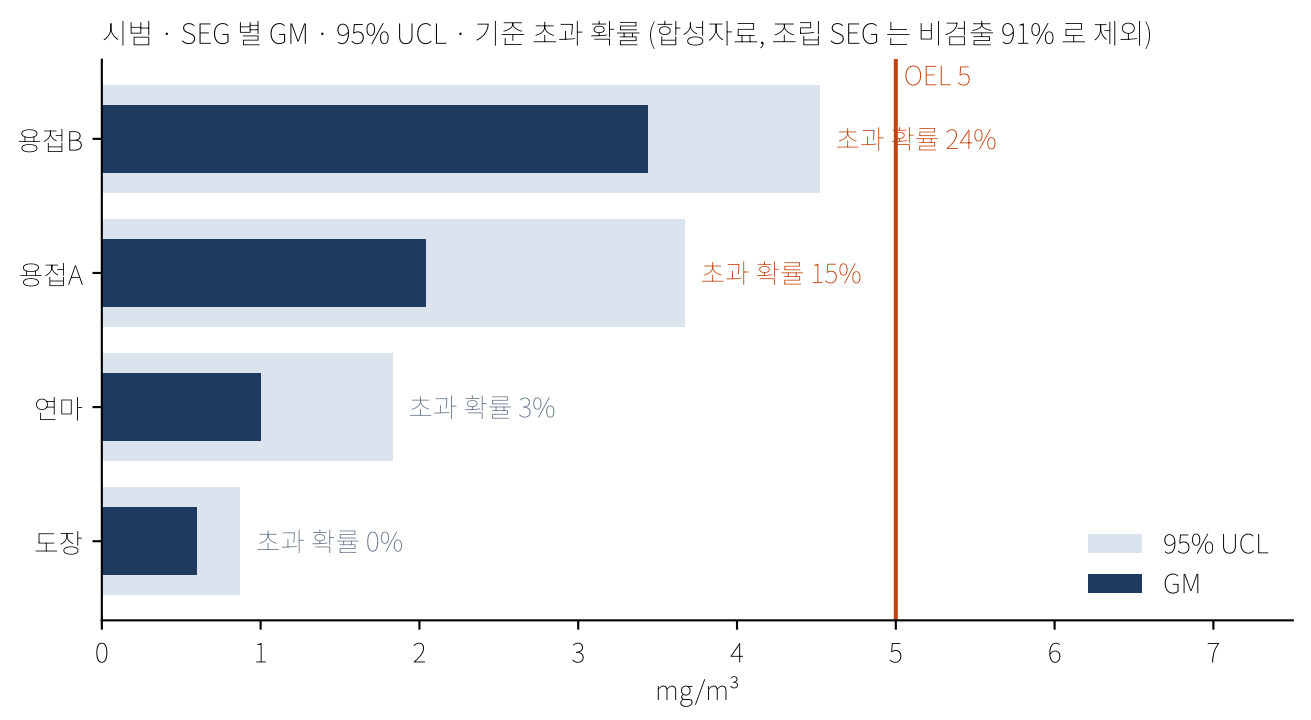
3. 하는 법 단계별 — 3.5 정의서 → 코드

단위·LOD·기준값·SEG·시료 유형·채취 시간을 **정의서**에 적고, 코드는 정의서를 구현합니다. 단위가 열마다 다르면 코드는 묻지 않고 계산합니다.

4. 시범 — 끝까지 한 번

합성 측정자료(SEG 5개, 232행)를 정의서 → 정제(중복 2, 15분 시료 1, 단위 ug→mg 12행, 480 mg/m³ 1) → 분포 → 판정까지 돌린 결과입니다.

4. 시범 — 끝까지 한 번


시범 · SEG 별 GM · 95% UCL · 기준 초과 확률 (합성자료, 조립 SEG 는 비검출 91% 로 제외)
OEL 5
용접B
초과 확률 24%
용접A
초과 확률 15%
연마
초과 확률 3%
도장
초과 확률 0%
95% UCL
GM
0
1
2
3
4
5
6
7
mg/m³

4. 시범 — 끝까지 한 번

SEG	n	비검출	GM	GSD	95% UCL	P>OEL	판정
용접B	49	0%	3.44	1.70	4.52	**24%**	조치 필요
용접A	60	2%	2.04	2.37	3.67	**15%**	조치 필요
연마	45	7%	1.00	2.34	1.83	3%	허용
도장	39	15%	0.60	1.86	0.87	<1%	허용
조립	35	91%	—	—	—	낮음	모수 보고 안 함

같은 용접B 를 정규 가정으로 내면 평균 3.9, t 구간 UCL 4.4 — "기준 이하"로 읽힙니다. 단위를 통일하지 않으면 GM 5.06 — 판정이 뒤집힙니다. 방법 문장 — "SEG 별 농도는 대수정규 분포를 가정하여 GM·GSD 를 추정하였고, 비검출은 좌측 중도절단 최대우도법으로 처리하였다. 노출기준 초과 확률과 95% UCL(Land 근사)로 판정하였다."

5. 주의점과 흔한 오류

오류	무엇이 잘못인가	대신
산술평균 ± t 구간	분포 가정이 틀림, 초과 확률을 못 냄	GM·GSD·UCL(Land)·P>OEL
"평균 < 기준 → 안전"	넷 중 하루는 넘음	초과 확률로 판정
비검출 삭제 / 0	위/아래 편향	MLE·ROS, 80% 넘으면 보고 안 함
단위 혼재	코드가 묻지 않음. 판정 뒤집힘	정의서 · 단위 통일 로그
지역시료로 개인 판정	개인 노출이 아님	개인시료
전후 비교 = 효과	생산량·계절이 달라짐	비교 조건 표, 보정

5. 주의점과 흔한 오류

오류	무엇이 잘못인가	대신
배경 무시	작업 기여 과대	배경 시료
기준값을 AI 에게 물음	기억 의존, 개정 반영 안 됨	고시 원문
SEG 당 시료 3개로 UCL	n 무시	6~10 이상

6. AI 와 함께 — 위임 가능 과업과 검증

> 자율성 수준 L0(쓰지 않음) · L1(보조 — 사람이 전부 검토) · L2(초안 — 전면 재검토) · L3(위임 — 표본 검수) · L4(자동 — 테스트 통과 시 채택). 정의는 Part 0 §2.2, 약어는 `00_공통/용어집.md`.

과업	수준	무엇을 맡기나	틀리면 어떻게 알아채나
기준값 찾기	L2 (원문 접지)	고시 원문을 붙이고 "이 문서에서 물질 X 의 TWA 와 단위를 문장 인용과 함께"	원문 대조. 원문 없이 물으면 위임 불가
정의서 초안	L2	코드북에서 단위·LOD·시료 유형 정리	사람이 전 항목 확인
합성 측정자료 생성	L3	GM·GSD·비검출·단위 함정 명세 → 코드	큰 n 에서 GM·GSD 복원
분석 스크립트	L3	단위 통일 로그·분포 검사·GM·GSD·UCL·P>OEL, **정규 결과 병기**	자기 합성자료 복원 · R `NADA` / `EnvStats` 재현
함정 검사	L3	단위 열을 읽는가 · 비검출 처리 줄 인용 · UCL 방법 · 비검출 >80% 처리	인용 줄 직접 확인
몬테카를로	명세 L0 · 코		

6. AI 와 함께 — 위임 가능 과업과 검증

과업	수준	무엇을 맡기나	틀리면 어떻게 알아채나
SEG 정의 · 판정 · 조치	L0	—	사람

프롬프트 원문은 `AI실습/00_README.md` (P-D01~D30).

7. 실습 — 상황별 가상 측정자료

회차	하는 일	주는 것 (발판)
9주	정의서(단위·기준값 원문·SEG·LOD) → 분석 계획 커밋 → 자기 합성자료 복원 → 배포본 정제(단위 12행·중복·불가능값)	시나리오 자료만 (`scenarios.py`)
10주	`seg` — SEG 표(GM·GSD·UCL·P>OEL, 정규 비교, 비검출 SEG 처리) · `background` — 작업 vs 배경	—
11주	`prepost` — GM 비 + **비교 가능성 표** · `trend` — 월별 추세 + 자기상관 · 방법 사실 목록 → 문장화	—
제출	P4 결과 보고 IV	루브릭 `30_실습.md`

8. 되돌아보기

- 오늘 잡은 오류 — 단위 열을 언제 봤는가. 못 봤다면 판정이 어떻게 뒤집혔는가
- "개선 효과가 있었다"고 쓰기 전에 무엇을 확인했는가
- 다음 Part 로 가져갈 것: **정의서가 코드를 정한다** — AI 개발에서는 그것이 명세와 평가셋입니다

참고문헌

- Rappaport et al. (1991). Assessment of long-term exposures to toxic substances in air. doi:10.1093/annhyg/35.1.61
- Helsel et al. (2006). Fabricating data: how substituting values for nondetects can ruin results, and what can be done about it. doi:10.1016/j.chemosphere.2006.04.051
- Hewett et al. (2007). A comparison of several methods for analyzing censored data. doi:10.1093/annhyg/mem045

(실존 확인: `reference/refs_verified.csv`)

가져갈 다섯 줄

1. 노출평가는 **분포**를 재는 연구다. 평균이 아니라 GM·GSD·UCL·**초과 확률**
2. 측정 설계는 사람의 일 — SEG · 시료 수 · **배경 농도** · **비교 가능한 조건** · 채취 보정
3. 비검출은 규칙대로. 80% 를 넘으면 모수를 보고하지 않는다
4. 단위·LOD·기준값은 **정의서**가 정한다. 기준값은 고시 원문에서
5. 전후 비교는 같은 조건에서만 근거다. 추세는 자기상관을 본다