

무엇이 칼슘 섭취량에 영향을 줄까?

김광훈 · 민영욱

수원대학교 통계정보학과

1. 주제 선정 배경 및 목적

성장기 아동에게 충분한 칼슘 섭취가 중요하다는 것은 자명한 사실이다. 그럼에도 불구하고 칼슘은 성장기 아동의 식생활에서 가장 결핍되기 쉬운 영양소 중 하나라고 알려져 있다. 따라서 칼슘 섭취량에 영향을 주는 요인을 찾아 보고자 한다.

참고문헌 : 아동영양학 리포트(<http://www.happycampus.com/doc/11208680>)

2. 연구 수행 방법

칼슘영양 증진을 위한 영양교육의 기초자료로 활용하기 위해 표본조사를 실시한 김태영(2008)의 “학교 우유급식 참여여부에 따른 초등학생의 칼슘 섭취실태 비교” 자료를 제공받아 분석을 수행하였다.

청주 시내 소재 6개 학교의 초등학생을 대상으로 416명에게 조사를 실시하였으며, 분석 변수로는 24개 식품 섭취량과 성별, 학년, 운동시간, 자신이 생각하는 체형, 컴퓨터 및 TV 사용시간, 우유 종류, 즐겨 먹는 유제품 종류, 우유 섭취 여부에 대한 변수 8개를 선정하였다.

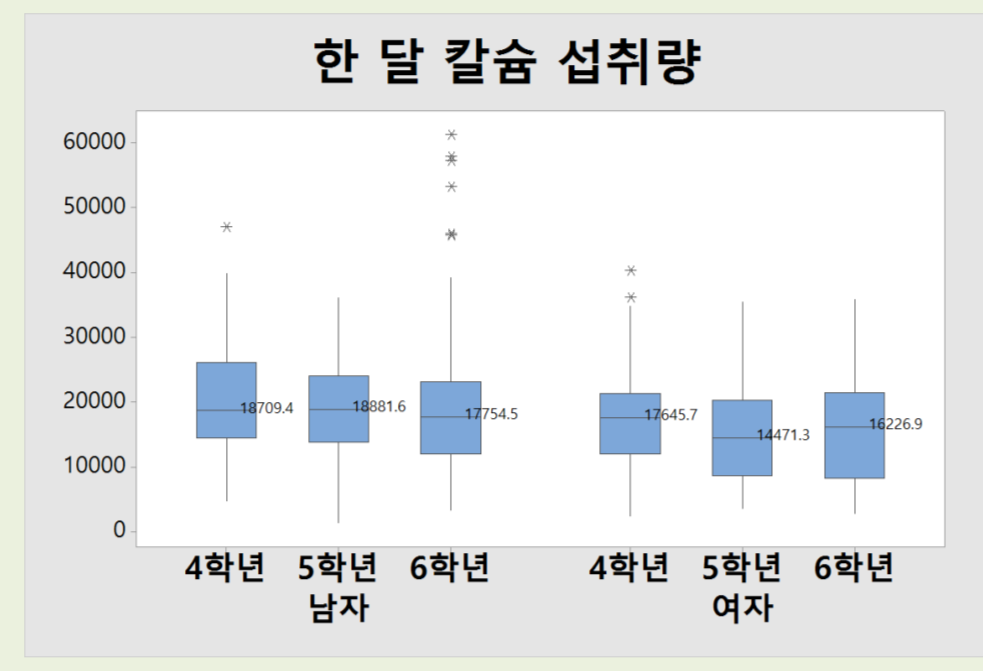
3. 분석 결과

1. 데이터 전처리

설문지 내에 초등학생의 칼슘 섭취실태를 조사하기 위해 선정된 24개 식품에 대해 섭취 횟수를 이용하여 각 학생별 한 달 칼슘 섭취량을 계산하였다.

한 달 칼슘 섭취량의 분포를 자세히 알아보기 위해 성별 학년별 한 달 칼슘 섭취량의 상자그림을 그려보았으며, 분포에 나타난 이상값 17개를 제거하였다. (그림 1)

출처 : 식품영양성분데이터베이스(www.foodnara.go.kr/kissna)



<그림 1. 성별 학년별 상자그림>

4. 평균 차이 검정

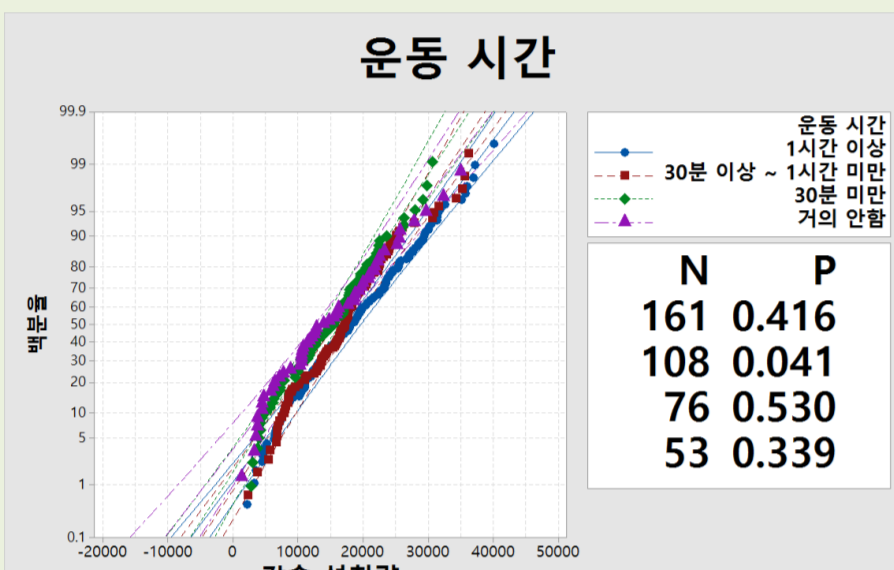
한 달 칼슘 섭취량의 평균이 운동 시간에 따라 차이가 있는지를 확인하기 위해 일원분산분석을 실시하였으며, p-값이 <0.001으로 매우 작으므로 집단 간 평균의 차이가 있다고 할 수 있다. 집단별 한 달 칼슘 섭취량의 평균 크기는 운동 시간이 많을수록 증가한다고 할 수 있다. (그림 7)

DF	Adj SS	Adj MS	F-값	P-값	
운동	3	911393067	303797689	5.26	0.001
오차	394	22770297230	57792633		
총계	397	23681690297			

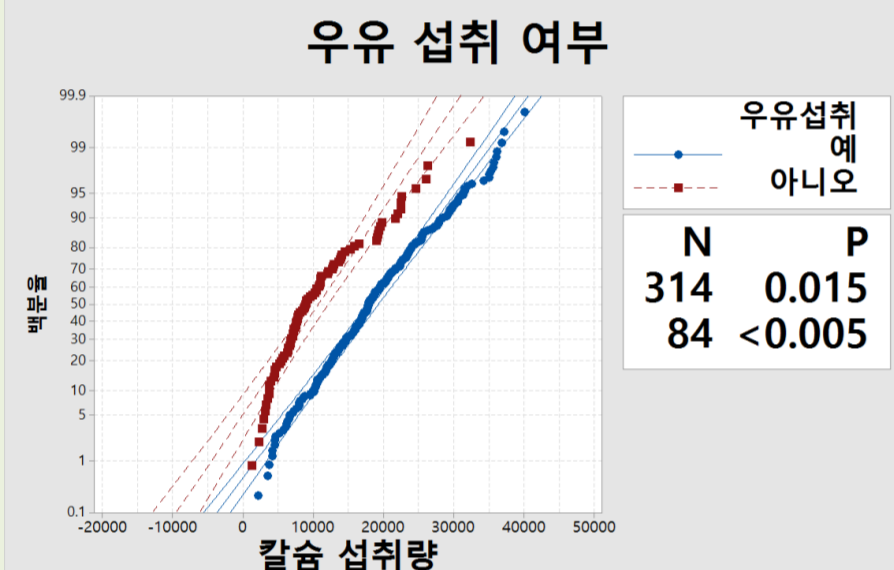
표준	N	평균	편차	95% CI
1시간 이상	161	18453	8058	(17275, 19631)
30분 ~ 1시간 미만	108	17028	7039	(15590, 18466)
30분 미만	76	14923	6941	(13209, 16638)
거의 안함	53	14842	8166	(12789, 16895)
합동 표준 편차				7602.15

<그림 7. 일원분산분석 결과>

2. 정규성 검정



<그림 2. 운동 시간별 정규 확률도>



<그림 3. 우유 섭취 여부에 따른 확률도>

운동 시간이 '1시간 이상', '30분 미만', '거의 안함' 데이터들은 p-값이 0.05 보다 크므로 정규성을 만족한다고 할 수 있다. 30분 이상 ~ 1시간 미만 데이터는 p-값이 0.041로 정규분포를 따른다고 할 수 없지만, N이 30보다 크므로 중심극한정리에 의해 정규성을 만족한다고 가정하였다. (그림 2)

우유 섭취 여부에 관한 데이터들은 '예'와 '아니오' 모두 p-값이 0.05보다 작으므로 규분포를 따른다고 할 수 없지만, N값이 30보다 크므로 중심극한정리에 의해 정규성을 만족한다고 가정하였다. (그림 3)

한 달 칼슘 섭취량의 평균이 우유 섭취 여부에 따라 차이가 있는지를 확인하기 위해 2-표본 t검정을 실시하였으며, p-값이 <0.001으로 매우 작으므로 두 집단 간 평균의 차이가 있다고 할 수 있다. 우유를 마시는 학생이 마시지 않는 학생보다 한 달 칼슘 섭취량의 평균이 7773mg 더 많다고 할 수 있다. (그림 6)

표준	N	평균	편차
예	314	18562	7173
아니오	84	10779	6561

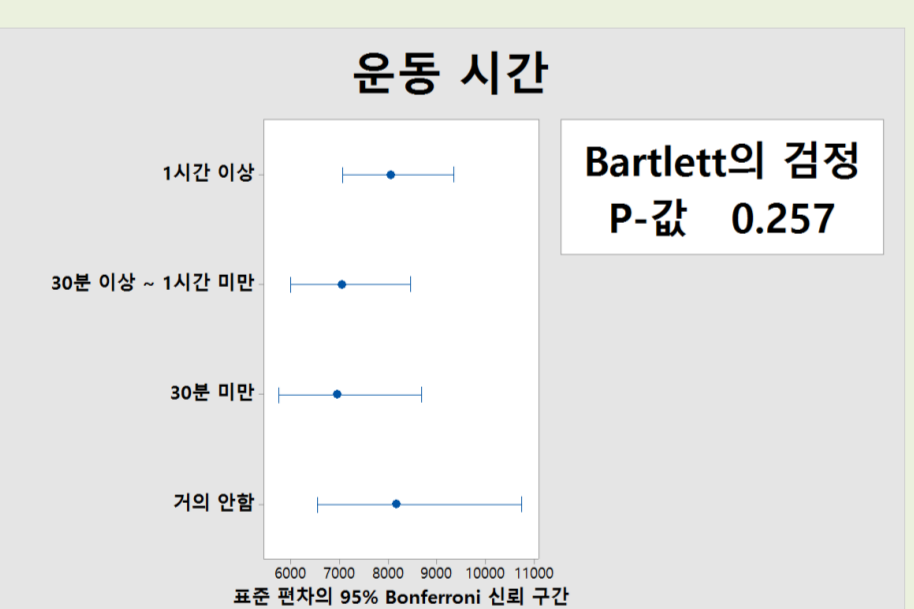
표준	95% CI
차이 추정치: 7773	(6070, 9475)
차이의 95% CI: (0 (대 #))	T-값 = 8.98, P-값 = 0.000, DF = 396

<그림 6. 2-표본 t검정 결과>

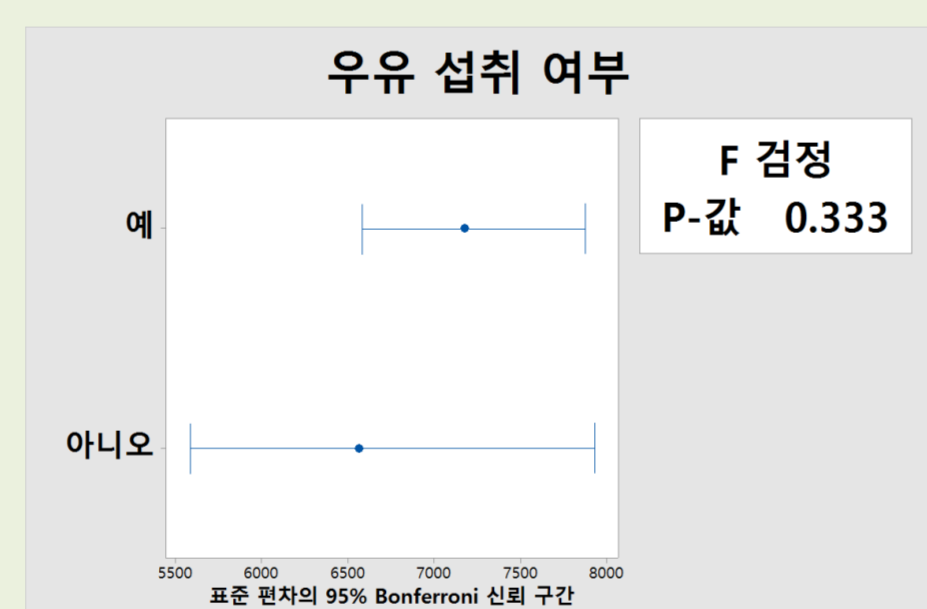
3. 등분산 검정

운동 시간 별로 한 달 칼슘 섭취량의 분산이 같은지를 확인하기 위하여 등분산 검정을 실시하였으며, Bartlett 검정 결과의 p-값이 0.257이므로 운동 시간별로 분산이 같다고 할 수 있다. (그림 4)

우유 섭취 여부에 따라 한 달 칼슘 섭취량의 분산이 같은지를 확인하기 위하여 등분산 검정을 실시하였으며, F 검정의 p-값이 0.333이므로 우유 섭취별로 분산이 같다고 할 수 있다. (그림 5)



<그림 4. 운동 시간별 등분산 검정>



<그림 5. 우유 섭취 여부에 따른 등분산 검정>

6. 다중 회귀 분석

단계적 회귀 분석

입력할 변수를 위한 $\alpha = 0.15$
제거할 변수를 위한 $\alpha = 0.15$

분산 분석

출처	DF	F-값	P-값
회귀	11	12.51	0.000
운동 시간	3	5.23	0.002
컴퓨터 TV 사용시간	3	2.65	0.048
우유종류	2	4.21	0.016
즐거 먹는 유제품 종류	2	4.00	0.019
우유섭취 여부	1	55.46	0.000

<그림 9. 회귀 분석 결과 1>

계수	계수	T-값	P-값	VIF
상수	18769	1896	11.07	0.000
운동 시간 / 1) 1시간 이상				
2) 30분 이상 ~ 1시간 미만	-619	-0.73	0.464	1.24
3) 30분 미만	-3383	-3.58	0.000	1.22
4) 거의 안함	-2823	-2.42	0.016	1.19
컴퓨터 TV 사용시간 / 1) 2시간 이상				
2) 1시간 이상 ~ 2시간 미만	1878	1.83	0.067	1.93
3) 1시간 미만	2576	2.53	0.012	1.95
4) 거의 안함	2748	2.50	0.013	1.76
우유종류 / 1) 흰 우유				
2) 가공 우유	-407	-0.43	0.665	1.36
3) 저지방 혹은 무지방 우유	-2220	-2.71	0.007	1.47
즐거 먹는 유제품 종류 / 1) 떠먹는 요구르트				
2) 마시는 요구르트	342	0.26	0.794	2.88
3) 치즈	-2263	-1.52	0.130	2.90
우유섭취 / 1) 예				
2) 아니오	-6455	-7.45	0.000	1.10

<그림 10. 회귀 분석 결과 2>

$$\text{한 달 칼슘 섭취량} = a + b_i \text{운동 시간}_i + c_j \text{컴퓨터 TV 사용시간}_j + d_k \text{우유 종류}_k + e_l \text{우유이외에 즐겨먹는 유제품종류}_l + f_m \text{우유 섭취 여부}_m, \quad i=1,2,3,4, \quad k=1,2,3, \quad l=1,2,3, \quad m=1,2$$

5. 다중 비교와 일반적 특성 한 달 칼슘 섭취량의 평균 비교

특성	N	평균 (±표준편차)	정규성	등분산성	평균비교
성별	남	188 18256 (±7625)	0.5333	0.996	<.001
	여	210 15708 (±7629)	0.011		
학년	4학년	119 18102 (±7456)	0.038	0.814	0.089
	5학년	141 16811 (±7833)	0.365		
	6학년	136 15937 (±7695)	0.157		
체형에 대한 자기 생각	비만	36 12742 (±7022)	0.091	0.304	0.007
	과체중	63 17459 (±6633)	0.247		
	정상	255 17173 (±7963)	0.122		
즐거 먹는 유제품 종류	떠먹는 요구르트	44 18023 (±7482)	0.844	0.965	<.001
	마시는 요구르트	30 16121 (±7894)	0.422		
	치즈	296 17858 (±7562)	0.206		
우유 종류	흰 우유	72 13351 (±7409)	0.062	0.313	<.001
	가공 우유	126 19134 (±7138)	0.231		
	저지방 우유 혹은 무지방 우유	90 18063 (±8217)	0.727		
컴퓨터 TV 사용 시간	2시간 이상	182 14803 (±7340)	0.023	0.587	<.001
	1시간 이상 ~ 2시간 미만	73 13710 (±6914)	0.053		
	1시간 미만	119 16688 (±7487)	0.270		
	거의 안함	123 18030 (±7704)	0.102		

<표 1. 일반적 특성별 한 달 칼슘 섭취량의 평균 비교>

다중 비교를 위해 Tukey의 모든 차이 비교 방법을 실시한 결과, '1시간 이상' 운동하는 학생과 '30분 미만' 운동하는 학생들의 한 달 칼슘 섭취량의 평균이 다르다고 할 수 있다. 운동 시간이 30분 이상인 두 집단 간 한 달 칼슘 섭취량의 평균이 같다고 할 수 있고, 1시간 미만인 세 집단 간 한 달 칼슘 섭취량이 같다고 할 수 있다. (그림 8)

나머지 분석 변수 6개에 대해 위와 같은 방법으로 분석을 실시하였고, 그 결과 값을 표 1에 정리하였다. (표 1)

Tukey 쌍별 비교				
Tukey의 방법 및 95% 신뢰 구간을 사용한 그룹화 정보				
N	평균	그룹화		
1시간 이상	161	18453	A	
30분 ~ 1시간 미만	108	17028	A B	
30분 미만	76	14923	B	
거의 안함	53	14842	B	

문자를 공유하지 않는 평균들은 유의하게 서로 다릅니다.

<그림 8. 다중 비교 결과>

4. 결론

일반량 분산 분석 결과, 성별, 학년, 운동시간, 자신이 생각하는 체형, 컴퓨터 및 TV 사용시간, 우유 종류, 즐겨 먹는 유제품 종류, 우유 섭취 여부가 각각 한 달 칼슘 섭취에 영향을 미치는 것으로 나타났다. 그러나 다중 회귀 분석 결과, 우유 섭취 여부가 가장 칼슘 섭취에 영향을 크게 미쳤으며, 그 다음으로는 운동 시간, 우유 종류, 즐겨 먹는 유제품의 종류, 컴퓨터 및 TV 사용시간 순으로 영향을 미친다고 결론 지을 수 있다.