In [1]:

*#* 연습문제 *4 p299, node (8)*

*#* 어떤 유형의 범죄발생건수가 대도시의 각 지역별로 다른지를 알아보기 위한 조사가 수행되었다*.* 특별히 조사대상으로 선정된

**from** scipy **import** stats

*#* 귀무가설*:* 범죄발생건수가 대도시의 각 지역과는 무관하다*.*

*#* 대립가설*:* 범죄발생건수가 대도시의 각 지역과는 무관하지 않다*.*

지역1 **=** [162, 118, 451, 18]

지역2 **=** [310, 196, 996, 25]

지역3 **=** [258, 193, 458, 10]

지역4 **=** [280, 175, 390, 19]

observed **=** [지역1, 지역2, 지역3, 지역4]

chi2\_stat, p\_value, dof, expected **=** stats**.**chi2\_contingency(observed)

print(f"검정통계량 : {chi2\_stat:.4f}") print(f"p-value : {p\_value}\n")

alpha **=** 0.01

**if** p\_value **<** alpha:

print(f"p-value는 {p\_value:.4f}로, 유의 수준 {alpha}보다 작다.\n따라서 귀무 가설을 기각한다.")

else:

print(f"p-value는 {p\_value:.4f}로, 유의 수준 {alpha}보다 크거나 같다.\n따라서 귀무 가설을 기각할 수 없음")

검정통계량 : 124.5297

p-value : 1.576242682023537e-22

p-value는 0.0000로, 유의 수준 0.01보다 작다.

따라서 귀무 가설을 기각한다.

In [3]:

*#* 연습문제 *2 p299, node (8) +* 데이터 시각화

**import** matplotlib.pyplot **as** plt

plt**.**rc('font', family**=**'Malgun Gothic') crimes **=** ['강간', '강도', '절도', '살인']

지역1 **=** [162, 118, 451, 18]

지역2 **=** [310, 196, 996, 25]

지역3 **=** [258, 193, 458, 10]

지역4 **=** [280, 175, 390, 19]

fig, ax **=** plt**.**subplots()

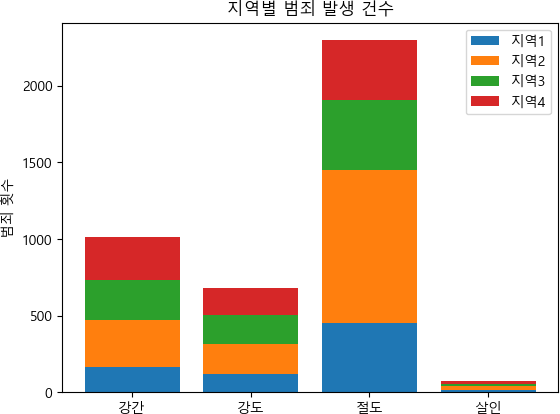
ax**.**bar(crimes, 지역1, label**=**'지역1')

ax**.**bar(crimes, 지역2, bottom**=**지역1, label**=**'지역2')

ax**.**bar(crimes, 지역3, bottom**=**[sum(x) **for** x **in** zip(지역1, 지역2)], label**=**'지역3') ax**.**bar(crimes, 지역4, bottom**=**[sum(x) **for** x **in** zip(지역1, 지역2, 지역3)], label**=**'지역4')

ax**.**set\_ylabel('범죄 횟수') ax**.**set\_title('지역별 범죄 발생 건수') ax**.**legend()

plt**.**show()



Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js