가설검정 및 구간추정에서 샘플크기 결정규칙의 고찰 및 유도

Review and Derivation of Sample Size Determination for Hypothesis Testing and Interval Estimation

최 성 운*
Sung-Woon Choi*

Abstract

Most useful statistical techniques in six sigma DMAIC are hypothesis testing and interval estimation. So this paper reviews and derives sample size formula by considering significance level, power of detectability and effect difference. The quality practioners can effectively interpret the practical and statistical significance with the rational sample sizing.

Keywords: DMAIC, Hypothesis Testing, Interval Estimation, Sample Size, Significance Level, Power of Detectability, Effect Difference

1. 서 론

식스시그마 DAMIC에서 치명인자를 선정하는 분석단계와 개선후 새로운 대체안(Alternative)을 비교하고 최적화하는 개선단계에서 가장 많이 사용되는 통계적 기법이 가설검정(Hypothesis Testing)과 구간추정(Interval Estimation)이다.

그러나 가설검정 및 구간추정은 Z, t, χ^2, F 분포모형의 가정, 샘플링 오차 α, β의 크기, 샘플의 크기 등에 따라 통계적 유의성(Statistical Significance)은 실무적 유의성(Practical Significance)과 전혀 다른 결론으로 나올 수 있다.[1]

* 가천대학교 산업공학과

따라서 본 연구에서는 DMIAC의 분석 및 개선단계에서 가장 유용하게 사용되는 가설검정 및 구간추정에서의 샘플크기 결정규칙을 고찰[2-5]하고 유도하기로 한다. 본 연구에서의 차별성은 식스시그마 DMAIC로 프로젝트를 수행하는 품질실무자가 목표효과차이, 유의수준, 신뢰수준, 검출력을 고려한 샘플크기의 결정규칙에 대한 적용과 이해를 돕는 데 있다.

2. 유의수준과 검출력을 고려한 샘플의 크기 결정

본 절에서는 가설검정에서 유의수준 α, 검출력 $1-\beta$, 효과차이(Effect Difference) δ를 고려한 샘플의 크기 n을 결정하는 규칙을 유도해본다. 유도된 공식은 좌측검정, 우측검정의 단측검정일 경우이고 α를 $\frac{\alpha}{2}$로 바꾸어 주면 양측검정(Not Equal의 H_1)에 대한 n의 규칙으로 활용할 수 있다.

2.1 1-Sample z Test

$$H_0 : \mu = \mu_0 \quad H_1 : \mu = \mu_1$$

$$\mu_0 + Z_\alpha \frac{\sigma}{\sqrt{n}} = \mu_1 - Z_\beta \frac{\sigma}{\sqrt{n}}$$

$$n = (\frac{(Z_\alpha + Z_\beta)\sigma}{\mu_1 - \mu_0})^2$$

2.2 1-Sample t Test

$$H_0 : \mu = \mu_0,\ H_1 : \mu = \mu_1$$

$$\mu_0 + t(n-1:\alpha)\frac{s}{\sqrt{n}} = \mu_1 - t(n-1:\alpha)\frac{s}{\sqrt{n}}$$

$$n = (\frac{(t(n-1:\alpha) + t(n-1:\beta))s}{\mu_1 - \mu_0})^2$$

2.3 2-Sample z Test

$$H_0 : \mu_1 - \mu_2 = 0,\ H_1 : \mu_1 - \mu_2 = \delta$$

$$0 + Z_\alpha \sqrt{\frac{2}{n}}\sigma = \delta - Z_\beta \sqrt{\frac{2}{n}}\sigma$$

$$n = 2(\frac{(Z_\alpha + Z_\beta)\sigma}{\delta})^2$$

2.4 2-Sample t Test

H_0 : $\mu_1 - \mu_2 = 0$, H_1 : $\mu_1 - \mu_2 = \delta$

$$0 + t(DF^*: \alpha)\frac{s_1}{\sqrt{n}} = \delta - t(DF^*: \beta)\frac{s_2}{\sqrt{n}}$$

$$n = (\frac{t(DF^*: \alpha)s_1 + t(DF^*: \beta)s_2}{\delta})^2$$

Where DF^*: Satterthwaite's Degree of Freedom(DF)

만약 등분산(Equal Variance)이라면 Pooled Variance

$s_p^2 = \dfrac{s_1 + s_2}{DF_1 + DF_2}$, $DF = DF_1 + DF_2 = 2n - 2$이고

$$n = 2(\frac{(t(DF: \alpha) + t(DF: \beta))s_p}{\delta})^2$$

2.5 Paired t Test

H_0 : $\mu_1 - \mu_2 = 0$, H_1 : $\mu_1 - \mu_2 = \delta$

$$0 + t(n-1: \alpha)\frac{s_d}{\sqrt{n}} = \delta - t(n-1: \beta)\frac{s_d}{\sqrt{n}}$$

$$n = (\frac{(t(n-1: \alpha) + t(n-1: \beta))s_d}{\delta})^2$$

2.6 1-Proportion Z Test

H_0 : $p = p_0$, H_1: $p = p_1$

$$p_0 + Z_\alpha\sqrt{\frac{p_0(1-p_0)}{n}} = p_1 - Z_\beta\sqrt{\frac{p_1(1-p_1)}{n}}$$

$$n = (\frac{Z_\alpha\sqrt{p_0(1-p_0)} + Z_\beta\sqrt{p_1(1-p_1)}}{p_0 - p_1})^2$$

2.7 2-Proportions Z Test

H_0 : $p_1 - p_2 = 0$, H_1: $p_1 - p_2 = \delta$

$$0 + Z_\alpha \sqrt{\frac{p_1(1-p_1)}{n} + \frac{p_2(1-p_2)}{n}} = \delta - Z_\beta \sqrt{\frac{p_1(1-p_1)}{n} + \frac{p_2(1-p_2)}{n}}$$

좌변은 H_0: $p_1 - p_2 = 0$에서 $p_1 = p_2 = p$, $p = \dfrac{r_1 + r_2}{n_1 + n_2}$이고

우변은 H_1: $p_1 - p_2 = \delta$에서 $p_2 = p_1 - \delta$이므로 각각을 대입하면

$$n = \left(\frac{Z_\alpha \sqrt{2p(1-p)} + Z_\beta \sqrt{p_1(1-p_1) + (p_1 - \delta)(1 - p_1 + \delta)}}{\delta}\right)^2$$

2.8 1-Sample Poisson Rate Z Test

H_0 : $u = u_0$, H_1: $u = u_1$

$$u_0 + Z_\alpha \sqrt{\frac{u_0}{n}} = \delta - Z_\beta \sqrt{\frac{u_1}{n}}$$

$$n = \left(\frac{Z_\alpha \sqrt{u_0} + Z_\beta \sqrt{u_1}}{u_1 - u_0}\right)^2$$

2.9 2-Sample Poisson Rates Z Test

H_0 : $u_1 - u_2 = 0$, H_1: $u_1 - u_2 = \delta$

$$0 + Z_\alpha \sqrt{\frac{u_1}{n} + \frac{u_2}{n}} = \delta - Z_\beta \sqrt{\frac{u_1}{n} + \frac{u_2}{n}}$$

좌변은 $H_0 : u_1 - u_2 = 0$에서 $u_1 = u_2 = u$, $u = \dfrac{c_1 + c_2}{n_1 + n_2}$이고

우변은 $H_1 : u_1 - u_2 = \delta$에서 $u_2 = u_1 - \delta$이므로 각각을 대입하면

$$n = \left(\frac{Z_\alpha \sqrt{2u} + Z_\beta \sqrt{u_1 + (u_1 - \delta)}}{\delta}\right)^2$$

2.10 1-Variance χ^2 Test

$$H_0 : \frac{\sigma^2}{\sigma_0} = 1, \qquad H_1 : \frac{\sigma^2}{\sigma_0^2} = \delta$$

$$\sigma_0^2 = \frac{(n-1)s^2}{\chi^2(n-1:\alpha)} \tag{1}$$

$$\sigma^2 = \frac{(n-1)s^2}{\chi^2(n-1:1-\beta)} \tag{2}$$

(2)식을 (1)으로 나누면

$$\delta = \frac{\chi^2(n-1:\alpha)}{\chi^2(n-1:1-\beta)} \tag{3}$$

(3)으로 n을 결정한다.

2.11 2-Variances F Test

$$H_0 : \frac{\sigma_1^2}{\sigma_2^2} = 1, \qquad H_1 : \frac{\sigma_1^2}{\sigma_2^2} = \delta$$

$$\frac{\sigma_1^2}{\sigma_2^2} = \frac{s_1^2}{s_2^2}\frac{1}{F(n-1,n-1:\alpha)} \tag{4}$$

$$\delta\frac{\sigma_1^2}{\sigma_2^2} = \frac{s_1^2}{s_2^2}\frac{1}{F(n-1,n-1:1-\beta)} \tag{5}$$

(5)식을 (4)으로 나누면

$$\delta = \frac{F(n-1,n-1:1-\beta)}{F(n-1,n-1:\alpha)} \tag{6}$$

(6) 식으로 n을 결정한다.

2.12 Correlation Coefficient Z Test

$$H_0 : \rho = \rho_0 \qquad H_1 : \rho = \rho_1$$

$$n = 2\left(\frac{Z_a + Z_\beta}{\frac{1}{2}ln\left(\frac{1+(\rho_1-\rho_0)}{1-(\rho_1-\rho_0)}\right)}\right)^2 - 3$$

2.13 One-Way ANOVA t Test

$$n = 2(\frac{\sigma}{E})^2 (t(k(n-1) : \frac{\alpha}{k(k-1)}) + t(k(n-1) : \beta))^2 \qquad (7)$$

Where k: # of Levels

n: # of Replicates

$k(n-1)$: DF of Error Term

$k(k-1)$: Bonferroni's ${}_kC_2$ Correction for Multiple Comparision Tests

초기에는 등분산 가정하에 t값을 Z값으로 사용하다가 Converge할 때까지 식(7)에 의해 n을 구한다.

2.14 2^k Factorial Design t Test

$$r = \frac{1}{2^{k-2}}(t(DF: \frac{\alpha}{2}) + t(DF: \beta)\frac{\sigma}{\delta})^2$$

Where r: # of Replicates

DF: DF of Error Term

σ: Standard Deviation of Error Term

3. 신뢰수준을 고려한 샘플크기 결정

본 절에서는 신뢰구간 추정에서 신뢰수준 $1-\alpha$, 오차마진(Margin of Error) E를 고려한 샘플의 크기 n을 결정하는 규칙을 유도해 본다. CI는 신뢰구간(Confidence Interval)을 의미한다.

3.1 1-Sample Z CI

$$\mu = x \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

$$E = \pm Z_{\frac{\alpha}{2}} \frac{\alpha}{\sqrt{n}}$$

$$n = (\frac{Z_{\frac{\alpha}{2}}\sigma}{E})^2$$

3.2 1-Sample t CI

$$\mu = \overline{x} \pm t(n-1:\alpha)\frac{s}{\sqrt{n}}$$

$$E = \pm t(n-1:\alpha)\frac{s}{\sqrt{n}}$$

$$n = (\frac{t(n-1:\alpha)s}{E})^2$$

3.3 2-Sample Z CI

$$\mu_1 - \mu_2 = (\overline{x_1} - \overline{x_2}) \pm Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}}$$

$$E = \pm Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}}$$

$$n = (\frac{Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}}}{E})^2$$

만약 $\sigma_1^2 = \sigma_2^2 = \sigma^2$ (Equal Variance)라면

$$n = (\frac{Z_{\frac{\alpha}{2}}\sqrt{2}\,\sigma}{E})^2$$

3.4 2-Sample t CI

$$\mu_1 - \mu_2 = (\overline{x_1} - \overline{x_2}) \pm t(DF^*:\alpha)\sqrt{\frac{s_1^2}{n} + \frac{s_2^2}{n}}$$

$$E = \pm (DF^*:\alpha)\sqrt{\frac{s_1^2}{n} + \frac{s_2^2}{n}}$$

$$n = (\frac{t(DF^*:\alpha)\sqrt{s_1^2 + s_2^2}}{E})^2$$

Where DF^*: Satterthwaite's DF

만약 $s_1^2 = s_2^2 = s_p^2$ (Pooled Variance)라면

$$n = (\frac{t(2n-2:\alpha)\sqrt{2}\,s}{E})^2$$

3.5 Paired t CI

$$\delta = \overline{d} \pm t(n-1:\alpha)\frac{s_d}{\sqrt{n}}$$
$$E = \pm t(n-1:\alpha)\frac{s_d}{\sqrt{n}}$$
$$n = (\frac{t(n-1:\alpha)s_d}{E})^2$$

3.6 1-Proportion Z CI

$$P = p \pm Z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}$$
$$E = \pm Z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}$$
$$n = (\frac{Z_{\frac{\alpha}{2}}\sqrt{p(1-p)}}{E})^2$$

3.7 2-Proportions Z CI

$$P_1 - P_2 = (p_1 - p_2) \pm Z_{\frac{\alpha}{2}}\sqrt{\frac{p_1(1-p_1)}{n} + \frac{p_2(1-p_2)}{n}}$$
$$E = \pm Z_{\frac{\alpha}{2}}\sqrt{\frac{p_1(1-p_1)}{n} + \frac{p_2(1-p_2)}{n}}$$
$$n = (\frac{Z_{\frac{\alpha}{2}}\sqrt{\frac{p_1(1-p_1)}{n} + \frac{p_2(1-p_2)}{n}}}{E})^2$$

만약 $p_1 = p_2 = p$라면

$$n = (\frac{Z_{\frac{\alpha}{2}}\sqrt{2p(1-p)}}{E})^2$$

3.8 1-Sample Poisson Rate Z CI

$$U = u \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{u}{n}}$$

$$E = \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{u}{n}}$$

$$n = \left(\frac{Z_{\frac{\alpha}{2}} \sqrt{u}}{E}\right)^2$$

3.9 2-Sample Poisson Rates Z CI

$$U_1 - U_2 = (u_1 - u_2) \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{u_1}{n} + \frac{u_2}{n}}$$

$$E = \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{u_1}{n} + \frac{u_2}{n}}$$

$$n = \left(\frac{Z_{\frac{\alpha}{2}} \sqrt{u_1 + u_2}}{E}\right)^2$$

만약 $u_1 = u_2 = u$라면

$$n = \left(\frac{Z_{\frac{\alpha}{2}} \sqrt{2u}}{E}\right)^2$$

3.10 Regression CI

3.10.1 β_1 CI

$$n = \frac{12}{k(k-1)(k+1)} \left(\frac{t(DF_E : \alpha)\sigma_E}{E \triangle x}\right)^2$$

Where k: # of x Regressor(Levels)

$$\triangle x = \frac{x_{\max} - x_{\min}}{k-1}$$

만약 2개의 양 극단의 x 수준에서 $(k=2)$에서 샘플의 크기는

$$n = 2\left(\frac{t(DF_E : \alpha)\sigma_E}{E \triangle x}\right)^2$$

3.10.2 β_0 CI

$$n = (\frac{t(DF_E : \alpha)\sigma_E^2}{E})^2 (1 + \frac{12}{k(k-1)(k+1)}(\frac{\overline{x}}{\triangle x})^2)$$

3.10.3 Predicted Value $\hat{y}$ CI

$$n = (\frac{t(DF_E : \alpha)\sigma_E^2}{E})^2 (1 + \frac{12}{k(k-1)(k+1)}(\frac{x - \overline{x}}{\triangle x})^2)$$

3.11 2^k Factional Design(FD) CI

$$r = \frac{1}{2^k}(\frac{t(DF_E : \alpha)\sigma_E}{E})^2$$

3.12 RSD(Response Surface Design) CI

3.12.1 ($2^K(FD)$ or $2^{k-p}(FFD)$ + Centers) Design

$$n = \frac{1}{2 \times 2^{k-1}}(\frac{t(DF_E : \alpha)\sigma_E}{E})^2$$

Where $DF_E = nN_{design} - 1 - DF_{Model}$

3.12.2 3^k FD

$$n = \frac{1}{2 \times 3^{k-1}}(\frac{t(DF_E : \alpha)\sigma_E}{E})^2$$

3.12.3 BBD(Box-Benhken Design)

$$n = \frac{1}{\frac{8}{3}\binom{k}{2}}(\frac{t(DF_E : \alpha)\sigma_E}{E})^2$$

3.12.4 CCD(Central Composite Design)

$$n = \frac{1}{2n_{star}^2 + n_{cube}}(\frac{t(DF_E : \alpha)\sigma_E}{E})^2$$

Where $N_{design} = n_{cube} + n_0 + n_{star}$

5. 결 론

본 연구에서는 DMAIC 분석단계 및 개선단계에서 가장 유용하게 사용되는 가설검정과 구간추정에서 샘플의 크기를 결정하는 규칙을 고찰하고 유형화하여 유도방안을 제시하였다.

가설검정에서는 유의수준, 검출력, 목표효과차이를 고려한 샘플크기 결정규칙을 14가지 유형으로 유도하였고, 구간추정에서는 신뢰수준, 오차마진을 고려한 샘플크기 결정규칙을 12가지 유형으로 제시하였다.

6. 참 고 문 헌

[1] 최성운(2012), “식스시그마 프로젝트에서 연구가설과 통계가설에 의한 통계적 유의성 및 실무적 유의성의 적용방안”, 대한안전경영과학회지, In Press.

[2] Chow S. Wang H. Shao J.(2007), Sample Size Calculations in Clinical Research, 2 Edition, Chapman and Hall/CRC.

[3] Julious S.A.(2008), Sample Sizes for Clinical Trials, Chapman and Hall/CRC.

[4] Machin D. et.al(2008), Sample Size Tables for Clinical Studies, 3 Edition, Wiley-Blackwell.

[5] Mathews P.G.(2005), Design of Experiments with MINITAB, ASQ.