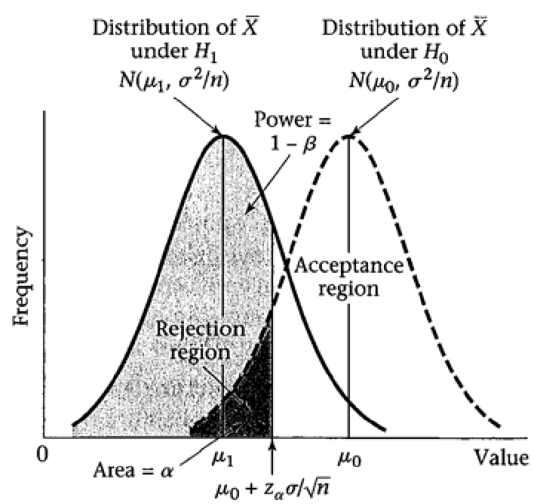
Hypothesis testing

February 11, 2014 Debdeep Pati

Power of a test

1. Assuming standard deviation is known. Calculate power based on one-sample z test. A new drug is proposed for people with high intraocular pressure (IOP), to prevent the development of glaucoma. A pilot study is conducted with the drug among 10 patients. Their mean IOP decreases by 5 mm Hg with a sd of 10 mm Hg after 1 month of using the drug. The investigator propose to study 100 participants in the main study. Is this a sufficient sample size for the study?
   $H_0 : \mu = \mu_0$ vs. $H_1 : \mu < \mu_0$ When the distribution is normal and variance is known. H_0 is rejected if $Z < Z_\alpha$ and H_0 is accepted otherwise. The best test does not depend on the alternative mean μ_1 What is the difference?
   Power = 1 - P(type II error) = P(reject H_0 | H_0 false) $= P(\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} < z_\alpha \mid \mu = \mu_1)$
   Under H_1, $\bar{X} \sim N(\mu_1, \sigma^2/n)$. Hence, Power $= P(\frac{\bar{X}-\mu_1}{\sigma/\sqrt{n}} < (\mu_0 + Z_\alpha \sigma/\sqrt{n} - \mu_1)/(\sigma/\sqrt{n})) =$ $P(Z < Z_\alpha + \frac{(\mu_0-\mu_1)\sigma}{\sqrt{n}})$.

Figure 7.5 **Illustration of power for the one-sample test for the mean of a normal distribution with known variance ($\mu_1 < \mu_0$)**



2. Compute the power of the test for the birthweight data with an alternative mean of 115 oz and $\alpha = 0.05$, assuming the true standard deviation = 24 oz. We have

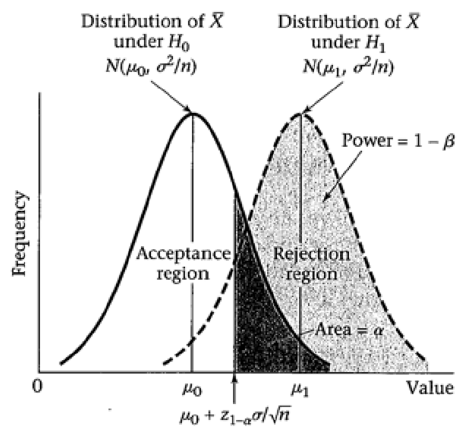
$\mu_0 = 120oz$, $\mu_1 = 115oz$, $\alpha = 0.05$, $\sigma = 24, n = 100$.
Power $= P(Z < Z_\alpha + \frac{(\mu_0-\mu_1)\sqrt{n}}{\sigma})$
$P(Z < -1.645 + 5(10)/24) = P(Z < 0.438) = 0.669$.

Alternative is greater than Null

1. The best test is H_0 rejected if $Z = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} > Z_{1-\alpha}$ and H_0 is accepted if $Z \leq Z_{1-\alpha}$. The power of the test is given by

$$\begin{aligned} P(\bar{X} > \mu_0 + Z_{1-\alpha}\sigma/\sqrt{n} \mid \mu = \mu_1) &= 1 - P(Z < Z_{1-\alpha} + \frac{(\mu_0 - \mu_1)\sqrt{n}}{\sigma}) \\ &= P(Z < Z_\alpha + \frac{(\mu_1 - \mu_0)\sqrt{n}}{\sigma}) \end{aligned}$$

Figure 7.6 Illustration of power for the one-sample test for the mean of a normal distribution with known variance ($\mu_1 > \mu_0$)



2. Using a 5% level of significance and a sample of size 10, compute the power of the test for the cholesterol data with an alternative mean of 190 mg/dL, a null mean of 175 mg/dL, and a standard deviation of 50 mg/dL. We have $\mu_0 = 175$, $\mu_1 = 190$, $\alpha = 0.05$, $\sigma = 50, n = 10$.

   Power $= P(Z < Z_\alpha + \frac{(\mu_0-\mu_1)\sqrt{n}}{\sigma})$
   $= P(Z < -1.645 + 15\sqrt{10}/50)) = P(Z < -0.696) = 0.243$

3. Hence power of a one sample Z test for the mean of a normal distribution with known variance (one-sided alternative) for the hypothesis

$$P(Z < Z_\alpha + \frac{(|\mu_0 - \mu_1|)\sqrt{n}}{\sigma})$$

4. Factors affecting the power:

   (a) If the significance level is made small (α decreases), Z_α decreases and hence the power decreases.

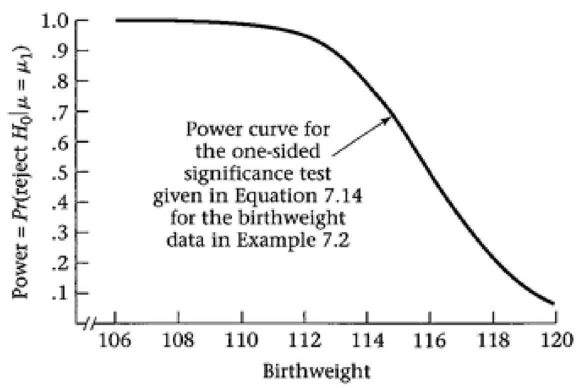
   (b) If the alternative mean is shifted further away from the null mean ($|\mu_0 - \mu_1|$ increases), then the power increases.

   (c) If the standard deviation of the distribution of individual observation increases (σ increases), then the power decreases.

   (d) If the sample size increases, then the power increases.

5. Power curve for the birthweight data:

Figure 7.7 **Power curve for the birthweight data in Example 7.2**



6. Two sided alternative: Reject if $\frac{(\bar{X}-\mu_0}{\sigma/\sqrt{n}} > Z_{1-\alpha/2}$ or $\frac{(\bar{X}-\mu_0}{\sigma/\sqrt{n}} < Z_{\alpha/2}$. Hence power

$$\Phi(Z_{\alpha/2} + \frac{(\mu_0 - \mu_1)\sqrt{n}}{\sigma}) + \Phi(Z_{\alpha/2} + \frac{(\mu_1 - \mu_0)\sqrt{n}}{\sigma})$$

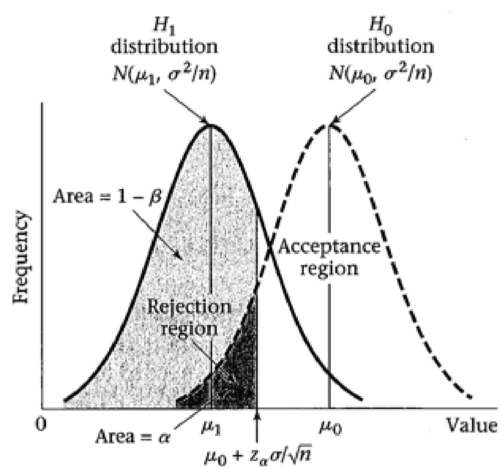
   and is approximated by

$$\Phi(Z_{\alpha/2} + \frac{|\mu_0 - \mu_1|\sqrt{n}}{\sigma})$$

1 Sample size determination

1. Sample size is important for a study design. Significance level is normally specified. If alternative hypothesis is true, what is the probability of rejecting null? This is the power of the test, typical values are 80% or above. Given a significance level α, and that the true alternative mean is expected to be μ_1, what sample size is needed to be able to detect a significance difference with probability $1-\beta$?

Figure 7.9 **Requirements for appropriate sample size**



2. For one sided alternative: Power $= P(Z < Z_\alpha + \frac{(|\mu_0-\mu_1|)\sqrt{n}}{\sigma}) = 1-\beta$. Solve n in terms of $\alpha, \beta, |\mu_0-\mu_1|$ and σ.

$$n = \frac{(Z_{1-\beta}+Z_{1-\alpha})^2\sigma^2}{(\mu_0-\mu_1)^2}$$

3. Consider the birthweight data. Suppose that $\mu_0 = 120$ oz, $\mu_1 = 115oz$, $\alpha = 0.05$, $\sigma = 24, 1-\beta = .80$, and we use a one-sided test. C ompute the appropriate sample size needed to conduct the test. $n = [242(Z_{0.8}+Z_{0.95})^2]/25 = 23.04(.84+1.645)^2 = 142.3$

4. For 2 sided alternative:

$$n = \frac{(Z_{1-\beta}+Z_{1-\alpha/2})^2\sigma^2}{(\mu_0-\mu_1)^2}$$

Factors affecting the sample size:

(a) The sample size increases as σ^2 increases

(b) The sample size increases as the significance level is made smaller (α decreases)

(c) The sample size increases as the required power increases ($1-\beta$ increases)

(d) The sample size decreases as the absolute value of the distance between the null and the alternative means $|\mu_0 - \mu_1|$ increases.

5. Consider a study of the effect of a calcium-channel-blocking agent on heart rate for patients with unstable angina. Suppose we want at least 80% power for detecting a significant difference if the effect of the drug is to change mean heart rate by 5 beats per minute over 48 hours in either direction and $\sigma = 10$ beats per minute. How many patients should be enrolled in such a study?

   We assume $\alpha = .05$ and $\sigma = 10$ beats. We use a two-sided test. $n = \frac{(Z_{1-\beta}+Z_{1-\alpha/2})^2\sigma^2}{(\mu_0-\mu_1)^2} = [100(Z_{0.8} + z_{0.975})^2]/25 = 31.36$. If we know the direction of the effect of the drug, $n = \frac{(Z_{1-\beta}+Z_{1-\alpha/2})^2\sigma^2}{(\mu_0-\mu_1)^2} = [100(Z_{0.8} + z_{0.95})^2]/25 = 24.7$.

Sample size calculation based on confidence interval width

We may want to estimate the effect with a given degree of precision. Suppose it is well known that propranolol lowers heart rate over 48 hours when given to patients with angina at standard dosage levels. A new study is proposed using a higher dose of propranolol than the standard one. Investigators are interested in estimating the drop in heart rate with high precision.

The $100(1-\alpha)\%$ CI for μ = true decline in the heart rate is $\bar{x} \pm t_{n-1,1-\alpha/2}s/\sqrt{n}$. The width of this CI is $2t_{n-1,1-\alpha/2}S/\sqrt{n}$. If we want the interval $< L$

$$2t_{n-1,1-\alpha/2}s/\sqrt{n} = L, \text{ or } n = 4t^2_{n-1,1-\alpha/2}S^2/L^2$$

Relationship between hypothesis testing and confidence intervals

Suppose we are testing $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$. H_0 is rejected with a two-sided level α test if and only if the two-sided $100(1-\alpha)\%$ confidence interval for μ does not contain μ_0. H_0 is accepted witha two-sided level α test if and only if the two-sided $100(1-\alpha)\%$ confidence interval for μ does contain μ_0. subsection*One sample χ^2 test for the variance or a normal distribution

1. An arteriosonde machine "prints" blood-pressure readings on a tape so that the measurement can be read rather than heard.

2. An argument for using such a machine is that the variability of measurement obtained by different observers on the same person will be lower than with a standard blood-pressure cuff.

3. Suppose we know from previously published work that $\sigma^2 = 35$ for d_i obtained from the standard cuff. We want to test the hypothesis

$$H_0 : \sigma^2 = \sigma_0^2 = 35 \, vs. \, \sigma^2 \neq \sigma_0^2$$

   How should we perform this test?

4.
$$\begin{aligned} X^2 &= \frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2_{n-1} \\ P(X^2 < \chi^2_{n-1,\alpha/2}) &= \alpha/2 = P(X^2 > \chi^2_{n-1,1-\alpha/2}) \end{aligned}$$

5. If $X^2 < \chi^2_{n-1,\alpha/2}$ or $X^2 > \chi^2_{n-1,1-\alpha/2}$, then H_0 is rejected.

6. $\chi^2_{n-1,\alpha/2} < X^2 < \chi^2_{n-1,1-\alpha/2}$, then H_0 is accepted.

Acceptance and rejection regions for the one-sample χ^2 test for the variance of a normal distribution (two-sided alternative)