I Sampling Distributions of Sample Variances.

① χ^2 : describes the distribution of sample variance

(1.1) Defn :-

$Z_i \sim N(0, 1)$; $i = 1, 2, \ldots, n$

$\chi^2(n) \sim \sum_{i=1}^{n} Z_i^2$; n is the no. of degrees of freedom

often written as $\chi^2(\nu)$; $\nu = n$

(1.2) Properties of χ^2 distribution

① χ^2 values ≥ 0.

② The shape of χ^2 distribution is different for different values of ν

③ Let $Y \sim \chi^2(\nu)$

$E(Y) = \nu$

$Var(Y) = 2\nu$.

(4) For $\nu \geq 30$; $\chi^2(\nu) \sim N(\nu, 2\nu)$ is a good approx^n!

i.e. $Z = \dfrac{\chi^2 - \nu}{\sqrt{2\nu}} \sim N(0, 1)$.

(1.3) Distribution of the Sample Variance.

This is a practical application of χ^2 distribution.

$Y_i \sim N(\mu, \sigma^2)$; $i = 1, 2, \ldots, n$

$\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$ where $S^2 = \dfrac{\sum(Y - \bar{Y})^2}{n-1}$ is the "unbiased" sample var.

Generated by CamScanner

2 important theorems

* Sampling D^n of the mean

the sampling D^n of $\bar{Y}$ from a random sample of size n drawn from a population w/ mean μ and variance σ^2 will have mean μ & variance $= \sigma^2/n$.

** Central Limit Th^m (CLT)

If random samples, Y_i of size n are taken from any D^n w/ mean μ & variance σ^2;

$$\bar{Y} \sim N(\mu, \sigma^2/n) \quad \text{as } n \to \infty.$$

(2) t Distribution

(2.1) Def^n

$$Z = \frac{\bar{Y} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

But often the population std. dev. σ is unknown.
If σ is replaced by the sample std. dev. S;
then Z is no more $N(0,1)$.

This led to the formulation of the t distribution by Gosset w/ ν degrees of freedom

$$t(\nu) = \frac{Z}{\sqrt{\frac{\chi^2(\nu)}{\nu}}} \quad ; \quad Z \sim N(0,1)$$

$\chi^2(\nu)$ is an independent χ^2 R.V. w/ ν deg. of freedom

(2.2) Application

Since $Z = \dfrac{\bar{Y} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$

and, $\chi^2(n-1) = \dfrac{(n-1)S^2}{\sigma^2}$ has χ^2 D^n w/ $n-1$ d.o.f.

$$T = \frac{Z}{\sqrt{\chi^2(n-1)/(n-1)}} = \frac{\bar{Y} - \mu}{S/\sqrt{n}} \sim t(n-1) \; ; \; \boxed{t(\infty) \sim N(0,1)}$$

Generated by CamScanner

F distribution (after Sir Ronald Fisher).

(3.1) Defⁿ

$$F(\nu_1, \nu_2) = \frac{\chi_1^2(\nu_1)/\nu_1}{\chi_2^2(\nu_2)/\nu_2}$$; χ_1^2 & χ_2^2 are independent of each other.

Also, if we have a sample of size n_i ; $i=1,2$ from a population w/ variance σ_i^2 ; $i=1,2$; each sample being independent of the other

$$F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}$$; S_i^2 is the variance estimate of σ_i^2 ; $i=1,2$

$\nu_i = (n_i - 1)$

(3.2) properties

→ i) F Dⁿ is defined for non-negative values.

ii) Not symmetric

---

④ Relationships among the Distributions

(i) $t(\infty) = Z \sim N(0,1)$

(ii) $Z^2 = \chi^2(1)$

(iii) $F(1, \nu_2) = t^2(\nu_2)$

(iv) $F(\nu_1, \infty) = \frac{\chi^2(\nu_1)}{\nu_1}$

Here ~ means "has the Dⁿ."

Generated by CamScanner