Sampling Distr

Suppose RV $\vec{X}=(X_1 \cdots X_n)$ form a random sample from a distr involving unknown param θ.

Let T be a function of $\vec{X}$ and possibly θ. That is, $T=r(\vec{X},\theta)$.

The distr of T (given θ) is called the sampling distr of T.

Purpose of Sampling Distr

If T is an estimator of θ (thus T is a function of $\vec{X}$ only)

We can use sampling distr of T to access how well T estimates θ.

Now denote T as $\hat{\theta}$,

$$\Pr(|\hat{\theta}-\theta|\le c)=\mathbb{E}_\theta\Big[\underbrace{\Pr(|\hat{\theta}-\theta|\le c\,|\,\theta)}_{\text{sampling distr}}\Big]$$

Chi-square (χ^2) Distr

It's a subfamily of Gamma distr, arising from sampling distr of variance estimators of random samples from normal.

For $m\in\mathbb{R}^+$, Gamma distr w/ $\alpha=m/2$, $\beta=1/2$ is called χ^2 distr w/ m degree of freedom.

- mean $=m$, var $=2m$
- MGF $\psi(t)=\left(\frac{1}{1-2t}\right)^{m/2}$ for $t<\frac{1}{2}$
- Additive property:

  If RV $X_1\cdots X_n$ independent and $X_i\sim\chi^2$ w/ m_i dof,

  RV $X=X_1+\cdots+X_n\sim\chi^2$ w/ $m_1+\cdots+m_n$ dof.

  (Follows from MGF)

Relating χ^2 and $N(0,1)$

If $X \sim N(0,1)$, RV $Y = X^2 \sim \chi^2$ w/ 1 dof

If $X_1 \cdots X_n \sim N(0,1)$, RV $X = \sum X_i^2 \sim \chi^2$ w/ n dof

Example: $X_1 \cdots X_n$ random sample from $N(0,1)$,

$\hat{\sigma}^2 = \frac{1}{n} \sum X_i^2$ which is a scaled χ^2 RV.

Joint Distr of Sample Mean $\hat{\mu}$ and Variance $\hat{\sigma}^2$

Suppose $X_1 \cdots X_n$ form a random sample from $N(\mu, \sigma^2)$.

Sample mean $\hat{\mu}$ and sample var $\hat{\sigma}^2$ are independent RVs.

$\bar{X}_n = \hat{\mu} \sim N(\mu, \frac{\sigma^2}{n})$, $\frac{n\hat{\sigma}^2}{\sigma^2} \sim \chi^2$ w/ n-1 dof

Intuition for χ^2: $\frac{n\hat{\sigma}^2}{\sigma^2} = \frac{\sum (X_i - \bar{X}_n)^2}{\sigma^2} = \sum \left(\frac{X_i - \bar{X}_n}{\sigma} \right)^2$

For $n \to \infty$, $\bar{X}_n \to \mu$, $\frac{X_i - \bar{X}_n}{\sigma} \to \frac{X_i - \mu}{\sigma} \sim N(0,1)$, we know $\left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi^2$

Now $\sum \left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi^2$ w/ n dof.

Substitute μ for it's estimator $\hat{\mu} = \bar{X}_n$, causing dof decrease by 1:

$\frac{n\hat{\sigma}^2}{\sigma^2} = \sum \left(\frac{X_i - \bar{X}_n}{\sigma} \right)^2 \sim \chi^2$ w/ n-1 dof.

(Proof refer to that of joint distr of LR $(\hat{\beta}, \hat{\sigma}^2)$ in Ch2)

- Sample mean and var are independent only if sampling from normal

Estimators of μ and σ^2

MLE of μ: $\hat{\mu} = \bar{X}_n$

MLE of σ^2: $\hat{\sigma} = \sqrt{\hat{\sigma}^2} = \sqrt{\frac{1}{n}\Sigma(X_i-\bar{X}_n)^2}$

An unbiased estimator of μ: $\mu' = \hat{\mu} = \bar{X}_n$

. . . σ^2: $\sigma' = \sqrt{\frac{1}{n-1}\Sigma(X_i-\bar{X})^2} = \sqrt{\frac{n}{n-1}}\,\hat{\sigma}$

Note, while the MLEs above are possibly unique to sampling from normal, the unbiased estimators are applicable to sampling from any distr.

probably approximately accurate

Usage Example: bounding sample size n s.t. $\hat{\mu}, \hat{\sigma}^2$ P.A.C.

Formulation: find n, s.t. $\Pr\left(|\hat{\mu}-\mu| \le \frac{1}{5}\sigma, |\hat{\sigma}-\sigma| \le \frac{1}{5}\sigma\right) \le \frac{1}{2}$

By independence of $\hat{\mu}, \hat{\sigma}$, $\Pr(\cdot,\cdot) = \overset{P_1}{\Pr(\cdot)} \cdot \overset{P_2}{\Pr(\cdot)}$

$P_1 = \Pr\left(\frac{\sqrt{n}|\hat{\mu}-\mu|}{\sigma} < \frac{1}{5}\sqrt{n}\right) = \Pr\left(|U| < \frac{1}{5}\sqrt{n}\right)$, $U \sim N(0,1)$

$P_2 = \Pr\left(0.8 \le \frac{\hat{\sigma}}{\sigma} \le 1.2\right) = \Pr\Big(0.64n \le \underbrace{\frac{n\hat{\sigma}^2}{\sigma^2}}_{\sim \chi^2 \text{ w/ } n-1 \text{ dof}} \le 1.44n\Big)$

t - distribution

For independent RV $Y \sim \chi^2$ w/ m dof, $Z \sim N(0,1)$,

RV $X = \dfrac{Z}{\sqrt{Y/m}}$ is t distributed.

- PDF: $f(x) = \dfrac{\Gamma(\frac{m+1}{2})}{\sqrt{m\pi}\,\Gamma(\frac{m}{2})} \left(1+\dfrac{x^2}{m}\right)^{-\frac{m+1}{2}}$ for $x \in (-\infty, \infty)$
- Mean: $\begin{cases} =0 & \text{for } m>1 \\ \text{not exist} & \text{for } m \le 1 \end{cases}$
- Var: $\begin{cases} \dfrac{m}{m-2} & \text{for } m>2 \\ \text{not exist} & \text{for } m \le 2 \end{cases}$

For $m>1$, k-th moment exists only for $k<m$.

Thus MGF doesn't exist.

t - distr and Random Sample from Normal

Suppose $X_1 \cdots X_n$ random sample from $N(\mu, \sigma^2)$

Let $\sigma' = \sqrt{\frac{1}{n-1}\sum (X_i - \bar{X})^2}$ (unbiased estimator of σ)

$\dfrac{\sqrt{n}}{\sigma'}(\bar{X}_n - \mu) \sim t$ distr w/ $n-1$ dof.

Proof: $\sigma'^2 = \dfrac{n}{n-1}\hat{\sigma}^2 = \underbrace{\boxed{\dfrac{n\hat{\sigma}^2}{\sigma^2}}}_{\chi^2 \text{ w/ } n-1 \text{ dof}} \cdot \dfrac{\sigma^2}{n-1} = \chi^2_{n-1} \cdot \dfrac{\sigma^2}{n-1}$

$\dfrac{\sqrt{n}}{\sigma'}(\bar{X}_n - \mu) = \dfrac{(\bar{X}_n - \mu)/(\sigma/\sqrt{n})}{\sigma'/\sigma}$ $\Rightarrow \sim N(0,1)$, as $\bar{X}_n \sim N(\mu, \frac{\sigma^2}{n})$

$= \dfrac{N(0,1)}{\sqrt{\chi^2_{n-1}\frac{\sigma^2}{n-1}}/\sigma} = \dfrac{N(0,1)}{\sqrt{\chi^2_{n-1}/(n-1)}} \sim t_{n-1}$

Note, if the denominator is $\sqrt{\hat{\sigma}^2/\sigma^2}$ instead of $\sigma'/\sigma = \sqrt{\sigma'^2/\sigma^2}$,

$$\hat{\sigma}^2/\sigma^2 = \frac{n\hat{\sigma}^2}{\sigma^2}/n = \chi^2_{n-1}/n$$

By substituting $\hat{\sigma}$ with σ, we make the factor match dof of χ^2 and thus getting a t distr.

Usage Example: make PAC arguments about sample mean $\hat{\mu}$.

Symbolically, $Pr(|\hat{\mu}-\mu| \le k\sigma') \ge$ threshold

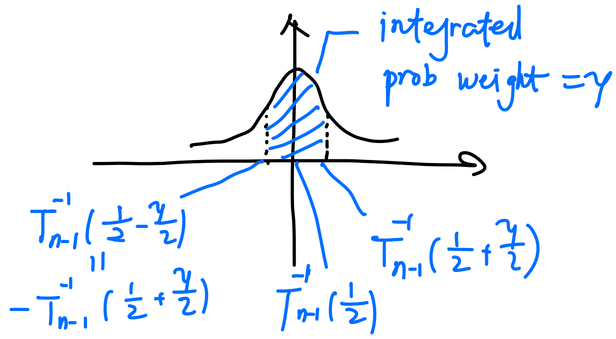
(confidence interval of μ given $\hat{\mu}$ calculated from observation)

As $U = \frac{\sqrt{n}}{\sigma'}(\hat{\mu}-\mu) \sim t_{n-1}$,

For confidence level γ,

a viable range of U is:

$$\left(-T^{-1}_{n-1}\left(\tfrac{1+\gamma}{2}\right),\ T^{-1}_{n-1}\left(\tfrac{1+\gamma}{2}\right)\right).$$



Distr of U

That translates to a γ-confidence interval of μ:

$$\left(\hat{\mu} - c\frac{\sigma'}{\sqrt{n}},\ \hat{\mu} + c\frac{\sigma'}{\sqrt{n}}\right) \text{ where } c = T^{-1}_{n-1}\left(\frac{1+\gamma}{2}\right)$$