Tutorial (For 09-22 and 09-25)

1. How to Compare two R.V. $\begin{cases} \mu_X \text{ v.s. } \mu_Y \to \mu_X - \mu_Y \\ \sigma_X^2 \text{ v.s. } \sigma_Y^2 \to \dfrac{\sigma_X^2}{\sigma_Y^2} \end{cases}$

- Assumption: ① X and Y are indepdent
  ② $X \sim N(\mu_X, \sigma_X^2)$, $Y \sim N(\mu_Y, \sigma_Y^2)$
  ③ $\sigma_X^2 = \sigma_Y^2 = \sigma^2$

- Point Guess (Point Estimation):

  Can use: $\bar{x}+\bar{y}$, $\bar{x}$, $\bar{y}$

  $\bar{x} - \bar{y} \to \mu_X - \mu_Y$ [BLUE]

- Interval Guess (Interval Estimation) and H.T.

Distribution of $(\bar{X} - \bar{Y})$?

Proof:

Method 1: get a rough idea:

$X \sim N(\mu_X, \sigma_X^2) \xrightarrow{\text{By Thm1(iii)}} \bar{X} \sim N(\mu_X, \sigma_X^2/n)$

$Y \sim N(\mu_Y, \sigma_Y^2) \to \bar{Y} \sim N(\mu_Y, \sigma_Y^2/m)$

$E(\bar{X}-\bar{Y}) = E(\bar{X}) - E(\bar{Y}) = \mu_X - \mu_Y$

$$Var(\bar{X}-\bar{Y}) = Var(\bar{X}) + Var(\bar{Y}) - 2Cov(\bar{X}, \bar{Y})$$
$$= \frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m} = \frac{\sigma^2}{n} + \frac{\sigma^2}{m}$$

the linear transform of 2 r.v.s that follow Normal will be still normal distributed: ($X-Y$, $X+Y$)

Method 2: (covered in Lecture (09-16)

$$X, Y \xrightarrow{\text{Thm1(iii)}} \bar{X}_n, \bar{Y}_m \xrightarrow{\text{Lemma 2}} \begin{pmatrix} \bar{X}_n \\ \bar{Y}_m \end{pmatrix}$$

$$\Big\downarrow \text{Lemma 3, } A = (1, -1)$$

distribution of $\bar{X}_n - \bar{Y}_n$.

Conclusion:

$$(\bar{X}_n - \bar{Y}_n) \sim N\left(\mu_X - \mu_Y, \frac{\sigma^2}{n} + \frac{\sigma^2}{m}\right) \quad \blacksquare$$

$$\frac{(\bar{X}_n - \bar{Y}_n) - (\mu_X - \mu_Y)}{\sigma\sqrt{\frac{1}{m} + \frac{1}{n}}} \sim \text{dist.}$$

What is σ^2 ?

☆ unbiased: If X is a r.v. and follow $N(\mu, \sigma^2)$ if $E[T(X)] = \mu$, the statistic $T(X)$ is unbiased.

$$\sigma^2 = \sigma_X^2 = \sigma_Y^2$$

$$E\left[\sum^n (X_i - \bar{X}_n)^2\right] = (n-1)\sigma^2$$

$$E\left[\sum^m (Y_j - \bar{Y}_m)^2\right] = (m-1)\sigma^2$$

$$\frac{\sum^n (X_i - \bar{X}_n)^2}{\sigma^2} \sim \chi^2_{(n-1)} \Rightarrow E\left[\frac{\sum^n (X_i - \bar{X}_n)^2}{\sigma^2}\right] \overset{\text{df}}{=} n-1. \quad ①$$

$$\frac{\sum^m (Y_i - \bar{Y}_m)^2}{\sigma^2} \sim \chi^2_{(m-1)} \Rightarrow E\left[\frac{\sum^m (Y_j - \bar{Y}_m)^2}{\sigma^2}\right] = m-1 \quad ②$$

$$\Rightarrow ① + ② = \underbrace{E\left[\frac{\sum^n (X_i - \bar{X}_n)^2}{\sigma^2} + \frac{\sum^m (Y_j - \bar{Y}_m)^2}{\sigma^2}\right] = (m+n-2)\sigma^2}_{S_p^2 =: \text{pooled sample Variance}}$$

$$\frac{(\bar{X}_n - \bar{Y}_m) - (\mu_X - \mu_Y)}{S_p \cdot \sqrt{\frac{1}{m} + \frac{1}{n}}}$$

$$(n+m-2)\, S_p^2 = (n-1)\, S_x^2 + (m-1)\, S_y^2$$

$$\Rightarrow \frac{(\bar{X}_n - \bar{Y}_m) - (\mu_X - \mu_Y)}{S_p \cdot \sqrt{\frac{1}{m}+\frac{1}{n}}} = \frac{\dfrac{(\bar{X}_n - \bar{Y}_n) - (\mu_X - \mu_Y)}{\sigma\sqrt{\frac{1}{n}+\frac{1}{m}}}}{\sqrt{\dfrac{(n+m-2)\, S_p^2/\sigma^2}{n+m-2}}} \Longleftrightarrow \boxed{\frac{N(0,1)}{\chi}}$$

$$\sim t_{n+m-2}$$

Conclusion:

$100(1-\alpha)\%$ C.I. for $\mu_X - \mu_Y$, when $\sigma_X^2 = \sigma_Y^2$ are unknown is. $\bar{x} + \bar{y} \pm t_{n+m-2, \frac{\alpha}{2}} \cdot S_p \sqrt{\frac{1}{n}+\frac{1}{m}}$.

exact value?

R language: qt, qf.

t-distribution critical value table

https://www.stat.tamu.edu/~lzhou/stat302/T-Table.pdf

2. Moment and Central moment

Definition (Population moments): For a positive integer k,

(i) the k^{th} population moment (about 0) of X is denoted by μ'_k and defined as

$$\mu'_k = E(X^k) \text{ if it exists.}$$

Note that μ'_1 is the population mean μ of X.

(i) the k^{th} population CENTRAL moment of X is denoted by μ_k and defined as

$$\mu_k = E(X-\mu)^k \text{ if it exists.}$$

Note that $\mu_1 = 0$. For population variance, we have $\boxed{\sigma^2 = E(X-\mu)^2 = \mu_2}$.

For the sake of convenience, we define $\sigma^k = (\sigma^2)^{k/2}$ for any positive integer k.

Caution: Except for $k = 2$, $\sigma^k \neq \mu_k$.

Name	Type	k.	explanation
Mean	Central	1	
Variance	Central	2	
Skewness	Central	3	symmetry
Kurtosis	Central	4	tail behavior

Theorem 2: Let X and Y be two random variables. Suppose that their mgfs, $M_X(t)$ and $M_Y(t)$, both exist and are equal for all t in $(-h, h)$ for some $h > 0$. Then, the distributions of X and Y are equal.

This theorem is particularly useful when the distribution of an independent sum of random variables has to be determined.

For instance, we can use it to prove Propositions 1 and 2, and (iii) of Theorem 1.

Noted:

If exist, $M_X(t) \Leftrightarrow F(X)$

assumption: Existence.

What if $M_X(t)$ doesn't exist ?

3. Moment Generating Function (MGF)

Definition (Moment generating function): The moment generating function (mgf) of a random variable X is denoted by $M_X(t)$ and is defined as $\boxed{M_X(t) = E(e^{tX})}$, if the expectation exists for t in a neighborhood of 0. To be more precise, there is a positive h such that, for all t in $(-h, h)$, $E(e^{tX})$ exists.

More explicitly, we write the mgf of X as

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x)\, dx$$

if the random variable X is continuous and $f(\cdot)$ is the pdf of X, or

$$M_X(t) = \sum_x e^{tx} p(x)$$

if X is discrete and $p(\cdot)$ is the pmf of X.

Distribution	Interpretation	E(X)	VAR(X)	M(t)=E(exp(tx))	Cdf(X) = P(X ≤ x) of Likelihood Func.	Pmf/Pdf
Binomial(n,p)	K successes in n Bernoulli trials	np	$np(1-p)$	$(1-p+pe^t)^n$	$L(\pi) = \binom{n}{k}\pi^k(1-\pi)^{n-k}$	$P(X=k) = \binom{n}{k}p^k q^{n-k}$
Bernoulli(p)	Probability of success	P	$p(1-p)$			$P(x) = p^x(1-p)^{1-x}$ if $x = 0$ or $x = 1$, 0 o.w.
Geom(p)	Prob that N trials for 1st success	$1/p$	$(1-p)/p^2$	$(e^t p)/(1-(1-p)e^t)$	$L(\pi) = (1-\pi)^n \pi$	$P(X=n) = p(1-p)^{n-1}$
Neg Bin(n,p)	Prob that N trials for R successes Generalization of Geometric **Sum of R independent geo RV's**	r/p	$r(1-p)/p^2$	$\left(\frac{e^t p}{1-(1-p)e^t}\right)^r$	$L(\pi) = \binom{N-1}{k-1}\pi^k(1-\pi)^{N-k}$	$P(X=k) = \binom{n-1}{r-1}p^r q^{n-r}$
Poisson(λ)	Limit of a binomial distribution as n→inf, p → 0. λ = rate per unit of time at which events occur. **Sum of Poi~Poi($\lambda 1+\lambda 2$)**	λ	λ	$e^{\lambda(e^t-1)}$	$L(\lambda) = \prod \frac{\lambda^{x_i}e^{-\lambda}}{x_i!}$	$P(X=k) = \frac{\lambda^k e^{-\lambda}}{k!}, k = 0,1,...$
N(μ, σ^2)	For X, Y ind., X~N(m1,v1), Y~N(m2,v2), then **X+Y~(m1+m2,v1+v2)**	μ	σ^2	$e^{\mu t}e^{\sigma^2 t^2/2}$	**No Closed Form for CDF** $L(\lambda) = \prod \frac{1}{\sqrt{2\pi}\sigma}\exp[\frac{x_i-\mu}{-2\sigma}]$	$\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$
Gamma(α, λ)	Sum of exponential RV's with parameter λ. **If sum of 2 exp RV, then α =2, and 2λ (if iid exp(λ))**	α/λ	α/λ^2	$\left(\frac{\lambda}{\lambda-t}\right)^\alpha, t < \lambda$		$\frac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}, x \geq 0$
Exp(λ)	Gamma with α = 1 **So if X~exp(λ), Y ~exp(λ), then X+Y ~ Gamma(2,2λ)**	$1/\lambda$	$1/\lambda^2$	$\lambda/(\lambda-t), t < \lambda$	$P(0 \leq X \leq x) = 1-e^{-\lambda x}$ for x ≥ 0, o o.w. → $P(X > x) = e^{-\lambda x}$ (x≥ 0)	$\lambda e^{-\lambda x}$ for x ≥ 0, 0 o.w.
Chi Sqr (n)	Gamma with a = ½, L = ½, n D.F.					
Uni[a,b]		(b+a)/2	$(b-a)^2/12$	$e^{\lambda(e^t-1)}$	x/(b-a) for x in [a,b], 0 o.w.	1/(b-a) for x in [a,b], 0 o.w.
Cauchy(θ, σ)	A special case of Student's T distribution, when d.f. = 1 (that is, X/Y for X, Y independent N(0,1)). **No Moments!**	Does Not Exist	Does Not Exist	Does Not Exist		$\frac{1}{\pi\sigma}\frac{1}{1+\left(\frac{x-\theta}{\sigma}\right)^2}$
Chi-Squared(p)	Sum of p iid Z^2 r.v., Z~N(0,1) Note: Sum of p independent X^2 is Chi-sq($df_1+...+df_p$)	P	2p	$(1-2t)^{-p/2}$		$\frac{(1/2)^{p/2}}{\Gamma(p/2)}x^{p/2-1}e^{-x/2}$

More details : http://www.stat.tamu.edu/~twehrly/611/distab.pdf

4. Central Limit Theorem (CLT)

> **Theorem 4 (Central Limit Theorem, standard version)**: Suppose that $\{X_n: n = 1,2,\dots\}$ is a sequence of i.i.d. random variables with a positive variance, each with an existing mgf $M_{X_n}(t)$, for all t in a neighborhood of 0. Denote by $\bar{X}_n$ the sample mean of $X_1, \dots, X_n$. Then, the limiting distribution of
>
> $$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \sim N(0, 1)$$
>
> is a standard normal distribution, where μ is the common population mean of $X_1, \dots, X_n$.

(Note that both μ and σ^2 are finite since the mgfs exist.)

- Why we need CLT ?

Before (in last tutorial),

We usually assume that a r.v. $X \sim N(\mu, \sigma^2)$.

this condition is not necessary and results are more general.

But indeed, if we don't have this assumption, the results can be valid.

- Population X is unknown.
- $\{X_n, n = 1, 2, 3 \dots\}$ is i.i.d. $\sigma_n > 0$
- If n is large enough

$$\Rightarrow \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \sim N(0, 1)$$

- Proof

Target: $\begin{cases} \dfrac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \sim N(0,1) \\ \text{By MGF} \end{cases}$

$$W_n := \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} = \frac{1}{\sqrt{n}} \cdot \sum_{i=1}^{n} V_i \quad \text{where } V_i := \frac{X_i - \mu}{\sigma^2}$$

$$M_{W_n}(t) = M_{\frac{1}{\sqrt{n}}\sum V_i}(t) = \prod_{i=1}^{n} M_{V_i}\left(\frac{t}{\sqrt{n}}\right)$$

$$M_{V_1}\left(\frac{t}{\sqrt{n}}\right) = E\left(e^{\frac{t}{\sqrt{n}} \cdot V_1}\right) = \sum_{i=0}^{\infty} \frac{\left(\frac{t}{\sqrt{n}}\right)^i}{i!} E(V_1^i)$$

$$= 1 + \frac{\frac{t}{\sqrt{n}}}{1!}\underbrace{\mu_1'}_{EX=0} + \frac{\left(\frac{t}{\sqrt{n}}\right)^2}{2!}\underbrace{\mu_2'}_{Var(X)=1.} + R_3\left(\frac{t}{\sqrt{n}}\right)$$

$$R_3\left(\frac{t}{\sqrt{n}}\right) = \sum_{k=3}^{\infty} \frac{\left(\frac{t}{\sqrt{n}}\right)^k}{k!} E(V_1^k)$$

If $r=2$, $g(\cdot) = M_{V_1}(\cdot)$, $a=0$, $x = \frac{t}{\sqrt{n}}$

$T_r(x) = 1 + \frac{1}{2}\frac{t^2}{\sqrt{n}}$, $\lim\limits_{n\to\infty} \dfrac{R_3\left(\frac{t}{\sqrt{n}}\right)}{\left(\frac{t}{\sqrt{n}}\right)^2} = 0$

(1) $t=0$ $\quad \lim\limits_{n\to\infty} [n R_3(0)] = 0$

(2) $t \neq 0$ $\quad \lim\limits_{n\to\infty} \left[n R_3\left(\frac{t}{\sqrt{n}}\right)\right] = \lim \left[t^2 \cdot \underbrace{\dfrac{R_3\left(\frac{t}{\sqrt{n}}\right)}{\left(\frac{t}{\sqrt{n}}\right)^2}}_{=0.}\right]$

$$\Rightarrow \lim_{n\to\infty} \left[n R_3\left(\frac{t}{\sqrt{n}}\right)\right] = 0 \quad \text{for all } t$$

$$\lim_{n\to\infty} M_{W_n}(t) = \lim_{n\to\infty} \left[1 + \frac{1}{2}\frac{t^2}{n} + R_3\left(\frac{t}{\sqrt{n}}\right)\right]^n$$

$$= \lim_{n\to\infty} \left[1 + \frac{\frac{t^2}{2} + nR_3\left(\frac{t}{\sqrt{n}}\right)}{n}\right]^n.$$

$$= e^{\lim\left[\frac{t^2}{2} + nR_3\left(\frac{t}{\sqrt{n}}\right)\right]} \longrightarrow e^{\frac{t^2}{2}} \longrightarrow \text{MGF of Normal}$$

Conclusion:

$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma}$ has a limiting distribution of $N(0,1)$ $\blacksquare$