Math 562 (Math. Stats)

Example 1:

$$\int \frac{1}{\theta} e^{-x/\theta}$$

$$\frac{1}{\theta}\int e^{-x/\theta}$$

1) $X \sim Exp(\theta)$

$$f(x) = \frac{1}{\theta} e^{-x/\theta}, \quad x > 0$$

and $Y = \ln X$.

Want to find pdf of Y.

Start by identifying domain: $Dom(Y) = (-\infty, \infty)$

$\ln(x) \in (-\infty, \infty), \ \forall x > 0$

Then for any $y \in (-\infty, \infty)$

$$P[Y \leq y] = P[\ln X \leq y]$$
$$= P[X \leq e^y]$$
$$= 1 - e^{(-1/\theta)(e^y)}$$

$$f_Y(y) = -e^{-1/\theta e^y}\left(-\frac{1}{\theta} e^y\right)$$
$$= \left(\frac{1}{\theta}\right) e^{-\frac{1}{\theta}e^y} \left(e^y\right)$$

Example 2:

If we have $X_1, \ldots, X_n \sim Exp(\theta)$, iid.

and $Y = \sum_{i=1}^{n} X_i$.

What is the distribution of Y?

(Rule of thumb: with sum of ind. variables, easiest to work with mgf)

$$M_Y(t) = E[e^{tY}] = E[e^{t\sum x_i}] = E[e^{tx_1 + tx_2 + \cdots}]$$
$$= E[e^{tx_1} e^{tx_2} \cdots e^{tx_n}] = E[e^{tx_1}]E[e^{tx_2}]\cdots E[e^{tx_n}]$$
$$= (E[e^{tx_1}])^n \quad \text{(because iid)}$$
$$= \frac{1}{(1-\theta t)^n} \sim GAM(\theta, n)$$

(See ☆$_2$)

For $X_1 \sim Exp(\theta)$

$$M(t) = E[e^{tX_1}] = \int_0^\infty e^{tx_1} f(x_1)\,dx$$

$$= \int_0^\infty e^{tx_1}\left(\frac{1}{\theta}\right)\left(e^{-\frac{x_1}{\theta}}\right)dx_1$$

☆2 (Derivation of MGF for Exponential distribution)

$$= \int_0^\infty \frac{1}{\theta} e^{-(1-\theta t)\frac{x_1}{\theta}}\,dx_1$$

$\left(s = \frac{x_1}{\theta},\ ds = \frac{1}{\theta}dx_1 \right)$

$$= \int_0^\infty e^{-(1-\theta t)s}\,ds$$

$$= \frac{1}{1-\theta t}\left[-e^{-(1-\theta t)s}\right]_0^\infty$$

$$= \frac{1}{1-\theta t}$$

What's Gamma?

$X \sim GAM(\alpha, \theta)$,

if $f(\alpha) = \frac{1}{\Gamma(\alpha)\theta^\alpha}(x^{\alpha-1})(e^{-\frac{x}{\theta}})$

$\Gamma(\alpha)$ Gamma function

$$M(t) = \frac{1}{(1-\theta t)^\alpha}$$

(when $\alpha = 1$, $Exp(\theta) = GAM(1, \theta)$)

$\Gamma(x) = \Gamma$

$$\chi^2(r) \sim GAM\left(\frac{r}{2}, 2\right)$$

$$M(t) = \frac{1}{(1-2t)^{r/2}}$$

$Y \sim GAM(\alpha, \theta)$

$X = \frac{2Y}{\theta} \sim \chi^2(2\alpha)$

$$E\left[e^{t\left(\frac{2Y}{\theta}\right)}\right] = \frac{1}{\left(1-\theta\left(\frac{2t}{\theta}\right)\right)^\alpha} = \frac{1}{(1-2t)^\alpha}$$

$\alpha = r/2$

$r = 2\alpha$

Example 3

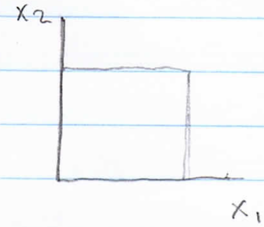
$X_1, X_2 \sim UNIF(0,1)$ and $X_1 \perp X_2$.

Consider $Y = X_1 + X_2$.

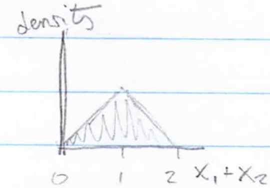
Determine the pdf of Y.

$Y \in (0,2)$. For $t \in (0,2)$,

$$P[Y \le t] = P[X_1 + X_2 \le t]$$



pdf of $X_1 + X_2$



imagine what happens as we increase n where we have $\sum_{i=1}^{n} X_i$

Example 4 Back to $N(\mu, \sigma^2)$

Take a random sample

$$X_1, X_2, \dots, X_n \sim N(\mu, \sigma^2)$$

Then $\bar{X} = \frac{1}{n}\sum X_i \sim N(\; , \;)$

$$M_{\bar{x}}(t) = E\left[e^{t\bar{x}}\right] = E\left[e^{(1/n)(t)(x_1)} \cdots e^{(\frac{1}{n})(t)(x_n)}\right]$$

$$= E\left[e^{(1/n)(t)(x_1)}\right] \cdots E\left[e^{(1/n)(t)(x_1)}\right]$$

$$= \left(E\left[e^{(\frac{1}{n})(t)(x_1)}\right]\right)^n$$

$$= \left[e^{\mu(\frac{t}{n}) + \frac{\sigma^2(\frac{t}{n})^2}{2}}\right]^n = e^{\left(\mu t + \frac{(\frac{\sigma^2}{n} \cdot t^2)}{2}\right)}$$

$$\sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

Let's Standardize!

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

$$\text{"}\lim\text{"}(W_n) = Z \sim N(0,1)$$

$\boxed{\text{The CLT}}$: Observe a $\overbrace{\text{random sample}}^{\text{this is shorthand for i.i.d.}}$ $X_1, X_2, \ldots, X_n$ from a population with mean μ and variance $\sigma^2 < \infty$.

Let $W_n = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{(\sum X_i) - (n\mu)}{(\sqrt{n})(\sigma)}$

Then $W_n \xrightarrow{d} Z \sim N(0,1)$ as $n \to \infty$.

Note

1) "$\xrightarrow{d}$" converges in dist. means that
$\lim_{n\to\infty} G_n(y) = G(y)$ for all y

(i.e. the sequence of function converges pointwise)

2) If $M_n(t) \to M(t)$ on $(-\delta, \delta)$,
then $W_n \xrightarrow{d} W$

☆ Q1 What's "lim"?

Q2 What if X_i are not normal?

Note: For any distribution, $E[\bar{x}]=\mu$, and $Var(\bar{x}) = \frac{\sigma^2}{n}$

where $\bar{x}$ is the mean of sample size n.

Hence $E[W_n]=0$, and $Var(W_n)=1$

" $\xrightarrow{d}$ "

1) $\lim\limits_{n\to\infty} G_n(x) = \Phi(x)$ for all x.

2) If $\lim\limits_{n\to\infty} M_n(t) = M(t)$ on $(-\delta, \delta)$, ← (for some $\delta>0$)

then $W_n \xrightarrow{d} W$.

this is sufficient, but not necessary

3) When n is large,

$W_n \approx Z$

$\bar{X}_n \approx N(\mu, \frac{\sigma^2}{n})$

(Since $W_n \xrightarrow{d} Z \sim N(0,1)$)

$\overline{\theta}$

$\boxed{\text{Example 1}}$

Let $X_1, X_2, \ldots, X_{100}$ be 100 observations from $Exp(1)$,

and $Y = \sum_{i=1}^{100} X_i$

$\frac{1}{(1-t)^n}$

Note: $Y \sim GAM(100, 1)$

Estimate $P[Y > 110]$.

$$Y = \sum_{i=1}^{100} x_i$$

Note: $\frac{110-100}{100} 10 = \frac{10}{100} \cdot 10 = \frac{100}{100} = 1$

$$P[Y > 110] = P\left[\frac{Y}{100} > \frac{110}{100}\right] = P\left[\underbrace{\frac{\frac{Y}{100} - 1}{1/\sqrt{100}}}_{\approx Z \text{ by CLT}} > \frac{\frac{110}{100} - 1}{1/\sqrt{100}}\right]$$

$$\approx P[Z > 1]$$
$$= 0.1587$$

Example 2

Let $X_1, X_2, \dots X_{15}$ be a random sample from the distribution with pdf

$$f(x) = \begin{cases} (3/2)\cdot x^2 & x \in (-1,1) \\ 0 & \text{otherwise} \end{cases}$$

Use CLT to estimate $P\left[-0.3 \le \sum_{i=1}^{15} X_i \le 1.5\right]$

$\mu = 0$

$$\sigma^2 = E[X^2] - E[x]^2$$

$$= \int_{-1}^{1} x^2 \left(\tfrac{3}{2}\right) x^2 \, dx - 0^2 = \int_{-1}^{1} \left(\tfrac{3}{2}\right) x^4 \, dx = \frac{3}{2}\left[\frac{x^5}{5}\right]_{-1}^{1} = \frac{3}{2}\left(\frac{1}{5} + \frac{1}{5}\right)$$

$$= \frac{3}{5}$$

(and so $\sigma = \sqrt{3/5} = \sqrt{0.6}$

So

$$P\left[-0.3 \le \sum_{i=1}^{15} X_i \le 1.5\right]$$

$$= P\left[\frac{-0.3}{\sqrt{15(0.6)}} \le \frac{\sum_{i=1}^{n} X_i - 0}{\sqrt{15 \cdot (0.6)}} \le \frac{1.5}{\sqrt{15 \cdot (0.6)}}\right]$$

(Recall $\dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \dfrac{\sum^{n} x_i - n\mu}{\sigma(\sqrt{n})}$

CLT

$$\approx P[-0.1 \le Z \le 0.5] = 0.6915 + 0.5398 - 1$$

$$= 0.2313$$

Sketch of the proof for CLT $\left(\begin{array}{c}\text{When we have} \\ M_n(t)\end{array}\right)$.

Consider $X_1, X_2, \ldots, X_n \sim Exp(\theta)$ [Note: $\mu = \theta = \sigma$]

$$W_n = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} = \frac{\sum^n X_i - n\mu}{\sigma(\sqrt{n})} = \frac{\sum X_i - n\theta}{(\sqrt{n})\theta}.$$

For each i, $X_i \sim Exp(\theta)$, $M_i(t) = (1-\theta t)^{-1}$

$$\sum_{i=1}^n X_i \sim M(t) = (1-\theta t)^{-n}$$

$$W_n = \frac{\sum X_i - n\theta}{(\sqrt{n})\theta}$$

$$M_{W_n}(t) = E\left[e^{tW_n}\right] = E\left[e^{t\left(\frac{\sum X_i - n\theta}{\sqrt{n}\theta}\right)}\right]$$

$$= E\left[e^{\frac{t\sum X_i}{\sqrt{n}\theta}} e^{-\frac{tn\theta}{\sqrt{n}\theta}}\right] = E\left[e^{\frac{t\sum X_i}{\sqrt{n}\theta}} e^{-t\sqrt{n}}\right]$$

$$= e^{-t\sqrt{n}} E\left[e^{\frac{t\sum X_i}{\sqrt{n}\theta}}\right] = e^{-t\sqrt{n}} \frac{1}{\left(1-\theta\cdot\left(\frac{t}{\sqrt{n}\theta}\right)\right)^n}$$

$$= e^{-t\sqrt{n}} \cdot \left(\frac{1}{1-\frac{t}{\sqrt{n}}}\right)^n = e^{-t\sqrt{n}}\left(1-\frac{t}{\sqrt{n}}\right)^{-n}$$

So $\ln M_{W_n}(t) = -\sqrt{n}\,t - n\ln\left(1-\frac{t}{\sqrt{n}}\right)$

$$= -\sqrt{n}\,t - n\left[-\left(\frac{t}{\sqrt{n}}\right) - \frac{\left(-\frac{t}{\sqrt{n}}\right)^2}{2} + \frac{\left(-t/\sqrt{n}\right)^3}{3} - \frac{\left(-t/\sqrt{n}\right)^4}{4} + \cdots\right]$$

$$= -\sqrt{n}\,t + \left[\sqrt{n}\,t + \frac{n\left(\frac{t^2}{n}\right)}{2} + \frac{n\left(\frac{t^3}{n(\sqrt{n})}\right)}{3} + \frac{n\left(\frac{t^4}{n^2}\right)}{4} + \cdots\right]$$

$\star \left(\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots \right)$

$$= 0 + \frac{t^2}{2} + \frac{1}{\sqrt{n}}\underbrace{\left[\frac{t^3}{3} + \frac{t^4}{(\sqrt{n})4} + \dots\right]}$$

this is a convergent series

Recall if n is large, then $\bar{X} \approx N(\mu, \sigma^2/\sqrt{n})$

Q So what if $n \to \infty$?

In some sense, $\bar{x}$ "converges" to the single point μ.

LLN Laws of Large Numbers

1. $\boxed{\text{Modes of convergence}}$

1) Convergence in distribution

$\lim_{n\to\infty} F_n(x) = F(x)$ for all points of continuity x for $F(x)$.

2) Convergence in probability

For any $\varepsilon > 0$,

$$\lim_{n\to\infty} P[\,|Y_n - Y| < \varepsilon\,] = 1$$

3) Convergence almost surely

$$P\left[\omega \,\middle|\, \lim_{n\to\infty} Y_n(\omega) = Y(\omega)\right] = 1$$

Weaker — Convergence in distribution ← (often called "weak convergence")

Strength:

Convergence in probability

Convergence almost surely

Stronger

Remark: convergence almost surely
⇒ convergence in probability
⇒ convergence in distribution

Def: $\boxed{\text{Stochastic Convergence}}$

$Y_n \longrightarrow c$ (stochastically) if $Y_n \xrightarrow{d} Y$ in probability where $P[Y=c]=1$.

(i.e. $\forall \varepsilon > 0 \ \lim_{n\to\infty} P[|Y_n - c| < \varepsilon] = 1$)

Equivalently, $\lim_{n\to\infty} M_n(t) = e^{ct}$ on $(-\delta, \delta)$ (for some $\delta > 0$)

Laws of Large Numbers

1) Bernoulli Law of Large Numbers

$$X_1, X_2, \ldots, X_n \sim BER(p) = BIN(1, p)$$

$$Y_n = \sum_{i=1}^{n} X_i \sim BIN(n, p)$$

$(q = 1-p)$

$$\overline{X} = \frac{1}{n} Y_n$$

$$M_{\overline{X}}(t) = E\left[e^{t \frac{1}{n} \sum x_i}\right] = \left(M\left(\tfrac{t}{n}\right)\right)^n = \left(pe^{\frac{t}{n}} + q\right)^n$$

$$= \left[p\left(1 + \frac{\left(\frac{t}{n}\right)^2}{2!} + \frac{\left(\frac{t}{n}\right)^3}{3!} + \ldots\right) + q\right]^n$$

$$= \left(1 + p\left(\tfrac{t}{n}\right) + p\frac{\left(\frac{t}{n}\right)^2}{2!} + \ldots\right)^n$$

$d(n) \to 0$ as $n \to \infty$

$$= \left[1 + p\left(\tfrac{t}{n}\right) + \frac{d(n)}{2}\right]^n$$

and

$$\left[1 + p\left(\tfrac{t}{n}\right) + \frac{d(n)}{2}\right]^n \longrightarrow e^{pt}$$

* (ask for this identity)

So $\overline{X} \xrightarrow{p} p$

Stochastically

$$\left[1 + \frac{z}{n}\right]^n \to e^z \quad \text{as } n \to \infty$$

$$\overline{X} \approx N\left(p, \frac{\sigma^2}{n}\right), \quad \frac{\sigma^2}{n} \to 0$$

2) Weak Law of Large Numbers

Let $X_1, X_2, \cdots, X_n$ be a random sample

from $f(x)$ with μ and $\sigma^2 < \infty$

Then $\bar{X} \xrightarrow{P} \mu$ (conv. in probability)

(i.e. $\lim_{n \to \infty} P\left[|\bar{X} - \mu| > \varepsilon \right] \le \lim_{n \to \infty} \frac{\sigma^2}{n \varepsilon^2} = 0$)

3) Strong Law of Large Number

If μ exists,

$$P\left[\omega \,\middle|\, \lim_{n \to \infty} \bar{X}(\omega) = \mu \right] = 1$$

i.e. $\bar{X} \to \mu$ almost surely

We will prove the weak LLN, but the strong LLN cannot be proven in this course (it's very difficult)

Proof of Weak LLN

(1) Let X be an RV., u a nonnegative function and $c > 0$.

Then
$$P[u(X) \geq c] \leq \frac{E[u(X)]}{c}$$

We only consider X has a pdf $f(x)$ (i.e. we ignore discrete case, whose proof is similar)

$$E[u(X)] = \int_{-\infty}^{\infty} u(x) f(x)\,dx = \int_{\{u(x) \geq c\}} u(x) f(x)\,dx + \int_{\{u(x) < c\}} u(x) f(x)\,dx$$

$$\geq \int_{\{u(x) \geq c\}} u(x) f(x)\,dx \geq c \int_{\{u(x) \geq c\}} f(x)\,dx$$

$$= c P[u(x) \geq c]$$

And so
$$P[u(x) \geq c] \leq \frac{E[u(x)]}{c}$$

(2) Markov Inequality

$u(x) = |X|^r$ and consider c^r in (1)

$$P\left[|X|^r \geq c^r\right] \leq \frac{E[|X|^r]}{c^r}$$

$$P\left[|X| \geq c\right] \leq \frac{E[|X|^r]}{c^r}$$

(3) Chebyshev inequality

$$P\left[|X-\mu| \geq k\sigma\right] \leq \frac{1}{k^2}$$

$u(x) = (x-\mu)^2$, $c = k^2\sigma^2$

$$P\left[(x-\mu)^2 \geq k^2\sigma^2\right] \leq \frac{E[(x-\mu)^2]}{k^2\sigma^2} = \frac{\cancel{\sigma^2}}{k^2\cancel{\sigma^2}}$$

$\hookrightarrow = P\left[|x-\mu| > k\sigma\right]$

Now $P\left[|\bar{X}-\mu| < \varepsilon\right] \geq 1 - \sigma^2/\varepsilon^2 n \xrightarrow[\text{as } n\to\infty]{} 1$

and $P\left[|X-\mu| < k\sigma\right] \geq 1 - 1/k^2$

$P\left[|X-\mu| < \varepsilon\right] \geq 1 - \frac{\sigma^2}{k^2}$ so $P\left[|X-\mu| \geq \varepsilon\right] = \frac{\sigma^2}{\varepsilon^2}$ Weak LLN

Example of application of CLT to Analysis

Show that

1). $\lim_{n\to\infty} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} = 1/2$

2) $\lim_{n\to\infty} \int_0^n \frac{e^{-t} t^{n-1}}{(n-1)!} dt = 1/2$

Proof of 1)

Let $X_1, X_2, \dots X_n$ be a random sample from $POI(1)$

and $Y_n = X_1 + X_2 + \dots + X_n \sim POI(n)$.

By CLT, $W_n = \frac{Y_n - n}{\sqrt{n}} \xrightarrow{d} Z \sim N(0,1)$

$P\left[\frac{Y_n - n}{\sqrt{n}} \le 0\right] = P\left[Y_n \le n\right]$ and this equals 1/2, which can be seen from CDF of Poisson

$= \lim_{n\to\infty} \sum_{k=0}^{n} \frac{e^{-n} n^k}{k!}$

Continued application of CLT to Analysis

Note:

$$\frac{\ln(n)}{n^r} \xrightarrow[n\to\infty]{} 0$$

Consider n^n, e^n, $n!$. Which grows faster?

There are faster.

So

$$\lim_{n\to\infty} \frac{n^n}{e^n n!} = 0$$

Prove it!

August 29

Recall from last class:

1) $$\lim_{n\to\infty} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} = 1/2.$$

by CLT

when $X_1, X_2, \dots, X_n \sim POI(1)$

$Y = \sum x_i \sim POI(n)$

$$P\left[\frac{Y-n}{\sqrt{n}} \le 0\right] = P[Y \le n] = \sum_{k=0}^{n} P[Y=k]$$

$$P[Z \le 0] = 1/2$$

$$= \sum_{k=0}^{n} \frac{e^{-n} n^k}{k!}$$

So

$$\lim_{n\to\infty} \sum_{k=0}^{\infty} \frac{e^{-n} n^k}{k!} = \lim_{n\to\infty} 1/2 = 1/2.$$

2) $$\lim_{n\to\infty} \int_0^n \frac{e^{-t} t^{n-1}}{(n-1)!}\, dt = 1/2$$

Let $X_1, X_2, \ldots, X_n \sim EXP(1)$

$$Y = \sum_{i=1}^{n} X_i \sim GAM(\underbrace{n}_{\alpha}, \underbrace{1}_{\theta})$$

$$f(t) = \frac{1}{\Gamma(n)\, 1^n} t^{n-1} e^{-t}$$

$$= \frac{1}{(n-1)!} t^{n-1} e^{-t}$$

So $P\left[\frac{Y-n}{\sqrt{n}} \le 0\right] = \lim_{n\to\infty} P[Y \le n] = \int_0^n \frac{e^{-t} t^{n-1}}{(n-1)!} dt$

$= P[Z \le 0]$

$= 1/2$

Taking the limit:

$$\lim_{n\to\infty} \int_0^n \frac{e^{-t} t^{n-1}}{(n-1)!} dt = \lim_{n\to\infty} 1/2 = 1/2$$

So we put 1) and 2) together

using IBP:

$u = \frac{t^{n-1}}{(n-1)!}$ $\quad dv = e^{-t} dt$

$du = \frac{t^{n-2}}{(n-2)!}$ $\quad -e^{-t}$

$$\int_0^n \frac{e^{-t} t^{n-1}}{(n-1)!} dt$$

$$= -e^{-t}\left(\frac{t^{n-1}}{(n-1)!}\right)\Big|_0^n + \int_0^n e^{-t} \frac{t^{n-2}}{(n-2)!} dt$$

$$= 1 + \int_0^n e^{-t}\left(\frac{t^{n-2}}{(n-2)!}\right) dt$$

$$= 1 - \sum_{k=0}^{n-1} \frac{e^{-n} n^k}{k!}$$

$$= 1 - \left(\sum_{k=0}^{n} \frac{e^{-n} n^k}{k!}\right) + \frac{e^{-n} n^n}{n!}$$

Take this limit:

Note:

$$\frac{1}{2} = 1 - \frac{1}{2} + \lim_{n\to\infty} \frac{e^{-n} n^n}{n!}$$

$$= \frac{1}{2} + \lim_{n\to\infty} \frac{e^{-n} n^n}{n!}$$

So

$$0 = \lim_{n\to\infty} \frac{e^{-n} n^n}{n!}$$

Sampling Distribution For $N(\mu, \sigma^2)$

Let $X_1, X_2, \ldots, X_n \sim N(\mu, \sigma^2)$.

1) $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

Then $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

$$M_{\bar{x}}(t) = E[e^{t\bar{x}}] = E[e^{\frac{t}{n}\sum X_i}] = E[e^{\frac{t}{n}X_1} e^{\frac{t}{n}X_2} \cdots e^{(\frac{t}{n})X_n}]$$

$$= \left(E[e^{(\frac{t}{n})X_1}]\right)^n = \left[e^{\mu(\frac{t}{n}) + \frac{\sigma^2(\frac{t}{n})^2}{2}}\right]^n$$

$$= e^{\mu t + \frac{\sigma^2 t^2}{2n}} = e^{\left(\mu t + \left(\frac{(\frac{\sigma^2}{n}) t^2}{2}\right)\right)} \sim N(\mu, \frac{\sigma^2}{n})$$

Special case of the "linear combination"

Example 1: Let $X_1, X_2, X_3 \sim N(\mu, 1/24)$ where $\mu \neq 0$ and indep.

Find a, b such that $Y = aX_1 + 4X_2 + bX_3 \sim N(0,1)$

$E[Y] = 0$

$E[Y] = a\mu + 4\mu + b\mu$

so

$a\mu + 4\mu + b\mu = 0$ and $\mu(a+b+4) = 0$ and

$a + b = -4$

$b = -(4+a)$ ☆$_1$

$Var(Y) = 1$

$Var(Y) = a^2\left(\frac{1}{24}\right) + 16\left(\frac{1}{24}\right) + b^2\left(\frac{1}{24}\right) = 1$

$a^2 + 16 + b^2 = 24$

$a^2 + b^2 = 8$ ☆$_2$

$a^2 + (4+a)^2 = 8$

$a^2 + a^2 + 8a + 16 = 8$

$2a^2 + 8a + 8 = 0$

$a^2 + 4a + 4 = 0$

$(a+2)^2 = 0$

So $a = -2, b = -6$

If $Z \sim N(0,1)$, then $Z^2 = V \sim \chi^2(1)$

— what we did at the beginning

If $\underbrace{Z_1, Z_2, \dots, Z_n} \sim N(0,1)$,

then $V = \sum_{i=1}^{n} Z_i^2 \sim \chi^2(n)$

! crucially! these are assumed independent.

$\boxed{\text{Third Question}}$ Sample Variance

$$S^2 = \frac{1}{(n-1)} \sum_{i=1}^{n} (x_i - x)^2$$

Sampling distribution

$\boxed{\text{Theorem}}$ Let $X_1, X_2, \dots, X_n$ be a random sample of size n from $N(\mu, \sigma^2)$. $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$

and $S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$.

Then:

1) $\bar{X}$ and S^2 are independent

2) $\frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{\sigma^2} \sim \chi^2(n-1)$

Proof (in 3 parts)

1) $$\underbrace{\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^2}_{\sim\chi^2(n)} = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\sigma^2} = \frac{\sum(x_i-\bar{x}+\bar{x}-\mu)^2}{\sigma^2}$$

by quadratic formula

$$= \sum\frac{(x_i-\bar{x})^2}{\sigma^2} + \sum\frac{(\bar{x}-\mu)^2}{\sigma^2} + 2\frac{\sum(x_i-\bar{x})(\bar{x}-\mu)}{\sigma^2}$$

$$= \sum\frac{(x_i-\bar{x})^2}{\sigma^2} + \frac{n(\bar{x}-\mu)^2}{\sigma^2} + \underbrace{2(\bar{x}-\mu)\frac{\sum(x_i-\bar{x})}{\sigma^2}}_{=0}$$

since $\sum x_i - n\bar{x} = 0$

$$= \underbrace{\frac{\sum(x_i-\bar{x})^2}{\sigma^2}}_{\text{we'll show } \sim\chi^2(n-1)} + \left(\frac{\bar{x}-\mu}{(\sigma/\sqrt{n})}\right)^2$$

Note: Class ended here

August 31

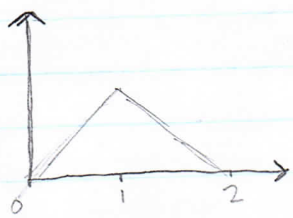
Note: Rule of thumb: When you're working with a sum of independent random variables, try to work with the MGF. Often clarifies things.

Let's talk about the sum of ind. unif. RVs.

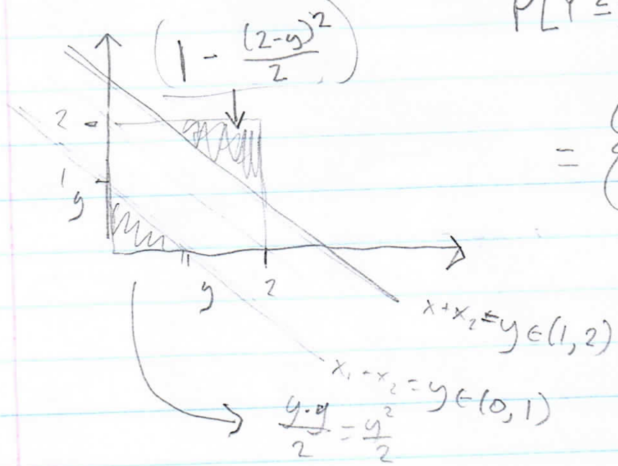
Consider $X_1, X_2 \sim UNIF(0,1)$ with $X_1 \perp X_2$.

$Y = X_1 + X_2$

Note the support of Y is $(0,2)$



$$f(y) = \begin{cases} y & y \in (0,1) \\ 2-y & y \in (1,2) \end{cases}$$



$$P[Y \le y] = P[X_1 + X_2 \le y]$$

$$= \begin{cases} \frac{y^2}{2} \\ \end{cases}$$

$$f(y) = \frac{\partial}{\partial y} P[Y \le y] = \frac{\partial}{\partial y} F(y)$$

$$f(y) = \begin{cases} y & y \in (0,1) \\ 2-y & y \in (1,2) \end{cases}$$

$$P[Y \ge 1] = 1 - P[Y < 1]$$

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad S^2 = \frac{1}{(n-1)}\sum (x_i - \bar{x})^2$$

Theorem

1) $\bar{X}$ and S^2 are independent

2) $$\frac{(n-1)S^2}{\sigma^2} = \frac{\sum (x_i - \bar{x})^2}{\sigma^2} \sim \chi^2(n-1)$$

If $X_1, X_2, \ldots, X_n$ are a random sample of size n from $N(\mu, \sigma^2)$

$\boxed{\text{Proof}}$

1) $$\underbrace{\sum \frac{(x_i-\mu)^2}{\sigma^2}}_{V \sim \chi^2(n)} = \underbrace{\sum \frac{(x_i-\bar{x})^2}{\sigma^2}}_{V_1 \sim ?} + \underbrace{\frac{n(\bar{x}-\mu)^2}{\sigma^2}}_{V_2 \ \chi^2(1)}$$

$$\Bigg(\underbrace{\frac{(x_i-\mu)}{\sigma}}_{Z}\Bigg)^2 \sim \chi^2(1)$$

If $\bar{X}$ and S^2 are independent, then we have:

$$\underbrace{M_V(t)}_{\chi^2(n)} = M_{V_1}(t)\underbrace{M_{V_2}(t)}_{\chi^2(1)}$$

$$\frac{1}{(1-2t)^{n/2}} = M_{V_1}(t)\left(\frac{1}{(1-2t)^{1/2}}\right)$$

So

$$\underbrace{M_{V_1}(t)}_{V_1 \sim \chi^2(n-1)} = \frac{(1-2t)^{1/2}}{(1-2t)^{n/2}} = \underbrace{\frac{1}{(1-2t)^{\frac{n-1}{2}}}}_{\text{Note this is the } \chi^2(n-1) \text{ mgf.}}$$

This shows that the second part of the theorem is true if the first part is true

Only need to show $\bar{X}$ and S^2 are indep.
In order to show that,
we only need to show X_i and S^2
are independent.

Second Step:
Show $\bar{X}$ and $(X_i - \bar{X})$ are independent.

For $X_1, X_2, \dots, X_n$ the joint density function is

$$f(x_1, x_2, \dots, x_n) = \frac{1}{(\sqrt{2\pi})^n \sigma^n} e^{\left[-\frac{1}{2\sigma^2}\sum (x_i-\mu)^2\right]}$$

$$= \frac{1}{(2\pi)^{n/2}(\sigma^2)^{n/2}} e^{\left[-\frac{1}{2\sigma^2}\left(\sum (x_i-\bar{x})^2 + n(\bar{x}-\mu)^2\right)\right]}$$

Transformation:

$y_1 = \bar{x}$, $y_i = x_i - \bar{x}$, $2 \le i \le n$

$$x_1 - \bar{x} = -\sum_{i=2}^{n} (x_i - \bar{x}) = -\sum_{i=2}^{n} y_i$$

$$\sum_{i=1}^{n} (x_i - \bar{x})^2 = \underbrace{\left(-\sum_{i=2}^{n} y_i\right)^2}_{(x_1-\bar{x})^2} + \underbrace{\sum_{i=2}^{n} y_i^2}_{\sum_{i=2}^{n}(x_i-\bar{x})^2}$$

$(x_i - \theta)$

$\alpha\theta^2$

The joint density for $Y_1 = \bar{X}$, $Y_i = X_i - \bar{X}$, $2 \le i \le n$

Jacobian, which is a constant since trans. is linear

$$g(y_1, y_2, \dots, y_n) = \frac{K}{(2\pi\sigma)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}\left(\left(-\sum y_i\right)^2 + \sum_{i=2}^{n} (y_i)^2 + n(y_1 - \mu)^2\right)\right]$$

$$= \frac{K}{(2\pi\sigma)^{n/2}} \exp\left[-\frac{1}{2\sigma^2} n(y_1 - \mu)^2\right] \exp\left[-\frac{1}{2\sigma^2}(\quad)\right]$$

y_i $2 \le i \le n$

Remark:

Why take $S^2 = \frac{1}{(n-1)} \sum (x_i - \bar{x})^2$?

$(n-1)$?

Consider $E[S^2]$.

We know $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

Note if $V \sim \chi^2(r)$, then $E(V) = r$ and $Var(V) = 2r$

Recall if
$X \sim GAM(\alpha, \theta)$
$E[X] = \alpha\theta$
$Var(X) = \alpha\theta^2$
$X \sim \chi^2(r) = GAM(\frac{r}{2}, 2)$
$E[X] = r$
$Var(X) = 2r$

So $E[S^2] = E\left[\frac{\sigma^2}{n-1} \frac{n-1}{\sigma^2} S^2\right]$

$$= \frac{\sigma^2}{n-1} E\left[\left(\frac{n-1}{\sigma^2}\right) S^2\right] = \frac{\sigma^2}{n-1} E\left[\frac{1}{\sigma^2} \sum (x_i - \bar{x})^2\right]$$

$$= \frac{\sigma^2}{(n-1)} (n-1) = \sigma^2$$

↑ "unbiased"

$$Var(S^2) = Var\left(\frac{\sigma^2}{n-1}\frac{(n-1)}{\sigma^2}S^2\right)$$

$$= \frac{\sigma^4}{(n-1)^2} Var\left(\frac{n-1}{\sigma^2}S^2\right) = \frac{\sigma^4}{(n-1)^2} \cdot 2(n-1)$$

$$= \frac{2\sigma^4}{(n-1)}$$

So far, we have from our theorem that

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \longrightarrow \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim Z \sim N(0,1)$$

$$S^2 \text{ ——— } \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

Two problems:

1) If σ^2 is unknown, we have to use S^2 to replace σ^2

What will be

$$\frac{X-\mu}{S/\sqrt{n}} \sim ?$$

2) To compare σ_1^2 and σ_2^2 by using S_1^2 and S_2^2,

we do

$$\frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}} \sim ?$$

Sep. 4

1.

meaning: $\cap$

$$P[X \le x, Y \le y] \quad — \quad P[\omega \mid X(\omega) \le x, Y(\omega) \le y]$$

ω denotes the compound event

i.e.

$$P[(X \le x) \cap (Y \le y)]$$

$X_1, X_2, \ldots, X_{100} \sim POI(\lambda)$

$$\bar{X} = \frac{1}{100} \sum_{i=1}^{100} x_i$$

Recall $\sum_{i=1}^{100} x_i \sim POI(100\lambda)$

$$P[\bar{X} \ge 1] = 1 - P[\bar{X} < 1]$$

(since Poisson RV is in $\{0\} \cup \{N\}$)

$$= 1 - P[\bar{X} = 0]$$

$$= 1 - P\left[\frac{1}{100} \sum x_i = 0\right]$$

☆!

$$= 1 - P\left[\sum x_i = 0\right]$$

$$= 1 - e^{-100\lambda}$$

Interesting note:
It's very hard to identify the sampling distribution of $\bar{x}$ for a random sample from a uniform dist.
The CLT is useful in this case.

Rule of thumb
When you have a random sample of 100, you should see 100 somewhere in your solution

Important!
It's only kosher to multiply by 100 here because the RHS is equal to zero.
Doesn't generally work for any real number.

Note: For HW2, we need to note that independence is a crucial condition

Sep. 4

meaning: $\cap$

$$P[X \le x, Y \le y] \text{ — } P[\omega \mid X(\omega) \le x, Y(\omega) \le y]$$

ω denotes the compound event

i.e.

$$P[(X \le x) \cap (Y \le y)]$$

$$X_1, X_2, \ldots, X_{100} \sim POI(\lambda)$$

$$\bar{X} = \frac{1}{100} \sum_{i=1}^{100} x_i$$

Recall $\sum_{i=1}^{100} x_i \sim POI(100\lambda)$

Interesting note:
It's very hard to identify the sampling distribution of $\bar{x}$ for a random sample from a uniform dist.
The CLT is useful in this case.

Rule of thumb
When you have a random sample of 100, you should see 100 somewhere in your solution

$$P[\bar{X} \ge 1] = 1 - P[\bar{X} < 1]$$

(since Poisson RV is in $\{0\} \cup \{\mathbb{N}\}$)

$$= 1 - P[\bar{X} = 0]$$

$$= 1 - P[\tfrac{1}{100} \sum x_i = 0]$$

$$= 1 - P[\sum x_i = 0]$$

$$= 1 - e^{-100\lambda}$$

Important!
It's only kosher to multiply by 100 here because the RHS is equal to zero.
Doesn't generally work for any real number.

Note: For HW2, we need to note that independence is a crucial condition

Sampling from $N(\mu, \sigma^2)$

Question: what if σ^2 is unknown?

1. $\bar{X} \sim N(\mu, \sigma^2) \longrightarrow \left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right) \sim N(0,1)$

   (More generally, $L = \sum_{i=1}^{n} a_i X_i$

   $L \sim N(\Sigma a_i \mu_i, \Sigma a_i^2 \sigma_i^2)$)

2. If $V = Z^2$,

   then $V \sim \chi^2(1)$.

   If $V = \sum \left(\frac{x_i - \mu_i}{\sigma_i}\right)^2$ then $V \sim \chi^2(n)$ (because of independence).

3. $S^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2$

   $E[S^2] = \sigma^2$ (ie, S^2 is an unbiased estimator of S^2)

   $$\frac{(n-1)S^2}{\sigma^2} = \frac{\sum (x_i - \bar{x})^2}{\sigma^2} \sim \chi^2(n-1)$$

4. $\dfrac{\bar{X}-\mu}{S/\sqrt{n}}$ We want to identify the distribution of this statistic.

$$\frac{\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\dfrac{(n-1)S^2/(\sigma^2)}{(n-1)}}}$$

$\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

$(n-1)S^2/(\sigma^2)$ we know $\sim \chi^2(n-1)$

<u>Def</u> (Student's t-distribution)

Let $Z \sim N(0,1)$, $V \sim \chi^2(r)$, <u>independent</u> (ie. $Z \perp V$)

this is very important

and $T = \dfrac{Z}{\sqrt{V/r}}$

We call t a Student's t-distribution, with degree of Freedom r denoted $t(r)$

To find pdf of this distribution:

Range of t: $(-\infty,\infty)$, Range of V: $(0,\infty)$

Joint Density of (Z,V)

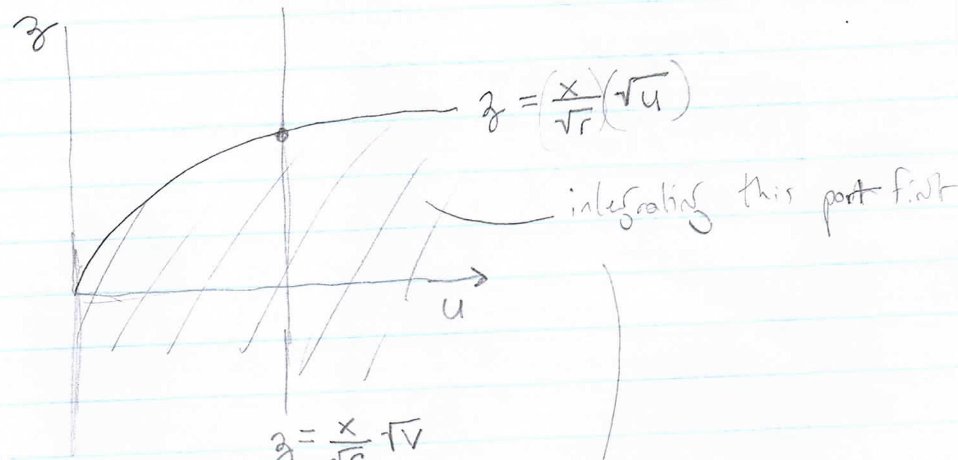
$$g(z,u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \frac{1}{\Gamma(\frac{r}{2})2^{r/2}} u^{\frac{r}{2}-1} e^{-\frac{u}{2}}$$

$z \in (-\infty,\infty)$, $u \in (0,\infty)$

For any $x \in (-\infty, \infty)$,

$$P[T \le x] = P\left[\left(Z/\sqrt{V/r}\right) \le x\right] = P\left[Z \le \left(\sqrt{V/r}\right) x\right]$$

Fixed



$$= \iint_{\Delta} g(z, u)\, dz\, du$$

$$= \int_0^{\infty} \int_{-\infty}^{\left(\frac{x}{\sqrt{r}}\right)\left(\sqrt{u}\right)} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \frac{1}{\Gamma\left(\frac{r}{2}\right) 2^{r/2}} u^{(r/2)-1} e^{-\frac{u}{2}}\, dz\, du$$

$$= \int_0^{\infty} \frac{1}{\sqrt{\pi}\,\Gamma\left(\frac{r}{2}\right)} \left[\int_{-\infty}^{\frac{x}{\sqrt{r}}\sqrt{u}} e^{-\frac{z^2}{2}}\, dz\right] \frac{1}{2^{\left(\frac{r+1}{2}\right)}} u^{\left(\frac{r}{2}-1\right)} e^{-\frac{u}{2}}\, du$$

And we differentiate w.r.t. x to get

$$f(x) = \frac{1}{\sqrt{\pi}\,\Gamma\left(\frac{r}{2}\right)} \int_0^{\infty} \left[e^{-\frac{\left(\frac{x}{\sqrt{r}}\sqrt{u}\right)^2}{2}} \frac{\sqrt{v}}{\sqrt{r}}\right] \frac{1}{2^{\left(\frac{r+1}{2}\right)}} u^{\frac{r}{2}-1} e^{-\frac{u}{2}}\, du$$

$$= \frac{1}{\sqrt{\pi}\,\Gamma(\frac{r}{2})} \int_0^\infty \frac{1}{2^{(\frac{r+1}{2})}\sqrt{r}}\, v^{\frac{r}{2}-1+\frac{1}{2}}\, e^{-\frac{v}{2}(1+\frac{x^2}{r})}\, dv$$

$$= \frac{1}{\sqrt{\pi r}\,\Gamma(\frac{r}{2})} \int_0^\infty \frac{1}{2^{\frac{r+1}{2}}}\, v^{\frac{r+1}{2}-1}\, e^{-(1+\frac{x^2}{r})\frac{v}{2}}\, dv$$

$$y = \left(1+\frac{x^2}{r}\right)v$$
$$dy = \left(1+\frac{x^2}{r}\right)dv$$
$$\frac{1}{\left(1+\frac{x^2}{r}\right)}\,dy = dv$$

$$= \frac{1}{\sqrt{\pi r}\,\Gamma(\frac{r}{2})} \int_0^\infty \frac{1}{2^{(\frac{r+1}{2})}} \left(\frac{y}{1+\frac{x^2}{r}}\right)^{\frac{r+1}{2}-1} e^{-\frac{y}{2}} \left(\frac{1}{1+\frac{x^2}{r}}\right) dy$$

$$= \frac{1}{\sqrt{\pi r}\,\Gamma(\frac{r}{2})} \left(\frac{1}{1+\frac{x^2}{r}}\right)^{\frac{r+1}{2}} \int_0^\infty \frac{1}{2^{(\frac{r+1}{2})}}\, y^{(\frac{r+1}{2}-1)}\, e^{-\frac{y}{2}}\, dy$$

multiply by $\left(\frac{\Gamma(\frac{r+1}{2})}{\Gamma(\frac{r+1}{2})}\right)$

Note this looks very similar to $\chi^2(r+1)$ pdf

$$= \frac{1}{\sqrt{\pi r}\,\Gamma(\frac{r}{2})} \left(\frac{1}{1+\frac{x^2}{r}}\right)^{\frac{r+1}{2}} \Gamma\left(\frac{r+1}{2}\right) \underbrace{\int_0^\infty \frac{1}{2^{(\frac{r+1}{2})}\Gamma(\frac{r+1}{2})}\, y^{(\frac{r+1}{2}-1)}\, e^{-\frac{y}{2}}\, dy}_{=1}$$

pdf of $\chi^2(r+1)$

$$= \frac{\Gamma(\frac{r+1}{2})}{\Gamma(\frac{r}{2})\sqrt{\pi r}} \left(1+\frac{x^2}{r}\right)^{-\left(\frac{r+1}{2}\right)}$$

this is the PDF of $t(r)$

(yay!)

$$\frac{\bar{X}-\mu}{(S/\sqrt{n})} \rightarrow T = \frac{Z}{\sqrt{V/r}}$$

$Z \sim N(0,1)$
$V \sim \chi^2(r)$
$Z \perp V$

T has pdf

$$f(x) = \frac{\Gamma\left(\frac{r+1}{2}\right)}{\Gamma\left(\frac{r}{2}\right)\sqrt{\pi r}}\left(1+\frac{x^2}{r}\right)^{-\left(\frac{r+1}{2}\right)} \qquad x \in (-\infty, \infty)$$

If $r > 1$, $E[T] = 0$.

If $r > 2$, $\operatorname{var}(T) = \frac{r}{r-2}$

$$\lim_{r \to \infty} T = Z$$

When r is large, ? 30

$T \approx Z$

Last Part: F distribution

Motivation: We wanted to compare σ_1^2 and σ_2^2 by comparing S_1^2 and S_2^2

$$N(\underline{\mu_1}, \sigma^2) \quad \text{and} \quad N(\underline{\mu_2}, \sigma^2)$$

$$\sigma_1^2 = \sigma_2^2$$

$$W = \frac{\left(\frac{S_1^2}{\sigma_1^2}\right)}{\left(\frac{S_2^2}{\sigma_2^2}\right)} = \frac{\underbrace{\frac{(n_1-1)S_1^2}{\sigma_1^2}}_{\chi^2(n_1-1)} \Big/ (n_1-1)}{\underbrace{\frac{(n_2-1)S_2^2}{\sigma_2^2}}_{\chi^2(n_2-1)} \Big/ (n_2-1)}$$

In defining the F-dist, we start with two χ^2-distributions.

Def. Let $U \sim \chi^2(r_1)$, $V \sim \chi^2(r_2)$

$U \perp V$. Define $W = \frac{(U/r_1)}{(V/r_2)}$

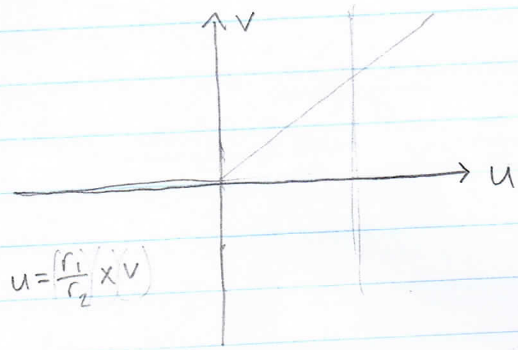
W is called $F(r_1, r_2)$ "F-distribution with r_1 and r_2 degrees of freedom"

Given U and V, the joint pdf is

$$g(u,v) = \frac{u^{\left(\frac{r_1}{2}-1\right)} e^{-\frac{u}{2}}}{\Gamma\left(\frac{r_1}{2}\right) 2^{(r_1/2)}} \; \frac{v^{\left(\frac{r_2}{2}-1\right)} e^{-\frac{v}{2}}}{\Gamma\left(\frac{r_2}{2}\right) 2^{(r_2/2)}}$$

For $x > 0$,

$$P[W \le x] = P\left[\frac{U/r_1}{(V/r_2)} \le x\right] = P\left[\frac{U}{r_1} \le x\left(\frac{V}{r_2}\right)\right]$$



$$= \int_0^\infty \int_0^{\left(r_1/r_2\right) v x} g(u,v)\, du\, dv$$

$$= \frac{1}{\Gamma\left(\frac{r_1}{2}\right)\Gamma\left(\frac{r_2}{2}\right)} \int_0^\infty \left[\int_0^{\left(\frac{r_1}{r_2} v x\right)} u^{\frac{r_1}{2}-1} e^{-\frac{u}{2}}\, du\right] \left(\frac{1}{2^{\frac{r_1+r_2}{2}}}\right) v^{\left(\frac{r_2}{2}-1\right)} e^{\left(-\frac{v}{2}\right)}\, dv$$

$$f(x) = \frac{1}{\Gamma\left(\frac{r_1}{2}\right)\Gamma\left(\frac{r_2}{2}\right)} \int_0^\infty \left(\frac{r_1}{r_2} v x\right) e^{\left(\frac{-\frac{r_1}{r_2} v x}{2}\right)} \left(\frac{r_1}{r_2} v\right) \left(\frac{1}{2^{\frac{r_1+r_2}{2}}}\right) v^{\left(\frac{r_2}{2}-1\right)} e^{\left(-\frac{v}{2}\right)}\, dv$$

See lecture notes on Blackboard

$$= \dots$$

$$= \frac{\Gamma\left(\frac{r_1+r_2}{2}\right)}{\Gamma\left(\frac{r_1}{2}\right)\Gamma\left(\frac{r_2}{2}\right)} \left(\frac{r_1}{r_2}\right)^{r_1/2} x^{\left(\frac{r_1}{2}-1\right)} \left(1 + \frac{r_1}{r_2} x\right)^{-\left(\frac{r_1+r_2}{2}\right)}$$

$\boxed{\text{Example 1}}$ Consider $X_1, X_2, X_3, X_4 \sim N(2,1)$
$Z_1, Z_2, Z_3, Z_4 \sim N(0,1)$
(independent)

Identify the sampling distribution of the following

1) $$\frac{\sqrt{2}(Z_1+Z_2)}{\sqrt{(X_1-X_2)^2+(Z_3+Z_4)^2}}$$

Note:
$Z_1+Z_2 \sim N(0,2)$
$Z_3+Z_4 \sim N(0,2)$
$X_1-X_2 \sim N(0,2)$
(These 3 variables are independent)

$$\frac{Z_1+Z_2}{\sqrt{2}} \sim N(0,1) \qquad \left(\frac{Z_3+Z_4}{\sqrt{2}}\right)^2 \sim \chi^2(1)$$

$$\frac{Z_3+Z_4}{\sqrt{2}} \sim N(0,1)$$

so

$$\left(\frac{X_1-X_2}{\sqrt{2}}\right)^2 \sim \chi^2(1)$$

$$\frac{X_1-X_2}{\sqrt{2}} \sim N(0,1)$$

and

$$\left(\frac{X_1-X_2}{\sqrt{2}}\right)^2+\left(\frac{Z_3+Z_4}{\sqrt{2}}\right)^2 \sim \chi(2)$$