Introductory Statistics

Tatiana Tyukina
tt51@leicester.ac.uk

Topic 2 - Estimators
2.4 - Inference based on other distributions
χ^2 *distribution*

- **Topic 0: Introduction**
- **Topic 1: Descriptive Statistics**
- **Topic 2: Estimators: Point estimation, Interval estimation**
  - 2.1 **Point estimation**,
  - 2.2 **Properties of estimators**,
  - 2.3 **Interval estimation**,
  - 2.4 Inference based on other distributions
- Topic 3: Hypothesis Testing
  - 3.1 Inference for two populations
  - 3.2 Hypothesis Testing Procedure
- Topic 4: Goodness of Fit: The χ^2 test
- Topic 5: Bayesian Estimation (optional)

INTERVAL ESTIMATION: RECAP

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

$$T\left(\frac{X}{n}, p\right) = \frac{X/n - p}{\sqrt{(X/n)(1-(X/n))/n}}, \quad T\left(\frac{X}{n}, p\right) \sim N(0,1)$$

$$T\left(\frac{X}{n}, p\right) = \frac{X/n - p}{\sqrt{1/4n}}, \quad T\left(\frac{X}{n}, p\right) \sim N(0,1)$$

$$T\left(\frac{X}{n}, p\right) = \frac{X/n - p}{\sqrt{p_g(1-p_g)/n}}, \quad T\left(\frac{X}{n}, p\right) \sim N(0,1)$$

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

DISTRIBUTIONS RELATED TO NORMAL DISTRIBUTION

Three important distributions:

- Chi-square (χ^2) distribution;
- F distribution;
- Student's t distribution.

GAMMA DISTRIBUTION

DEFINITION

*The **gamma function**, denoted by $\Gamma(a)$, is defined as*

$$\Gamma(a) = \int_0^\infty e^{-x} x^{a-1} dx, \quad a > 0.$$

GAMMA DISTRIBUTION

DEFINITION

*The **gamma function**, denoted by $\Gamma(a)$, is defined as*

$$\Gamma(a) = \int_0^\infty e^{-x} x^{a-1} dx, \quad a > 0.$$

It can be shown that for $a > 0$, $\Gamma(a) = (a-1)\Gamma(a-1)$

GAMMA DISTRIBUTION

DEFINITION

*The **gamma function**, denoted by $\Gamma(a)$, is defined as*

$$\Gamma(a) = \int_0^\infty e^{-x} x^{a-1} dx, \quad a > 0.$$

It can be shown that for $a > 0$, $\Gamma(a) = (a-1)\Gamma(a-1)$

$$\Gamma(n) = (n-1)!$$

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

GAMMA DISTRIBUTION

DEFINITION

*The **gamma function**, denoted by $\Gamma(a)$, is defined as*

$$\Gamma(a) = \int_0^\infty e^{-x} x^{a-1} dx, \quad a > 0.$$

DEFINITION

*A random variable X is said to possess a **gamma probability distribution** with parameters $\alpha > 0$ and $\beta > 0$ if it has the pdf given by*

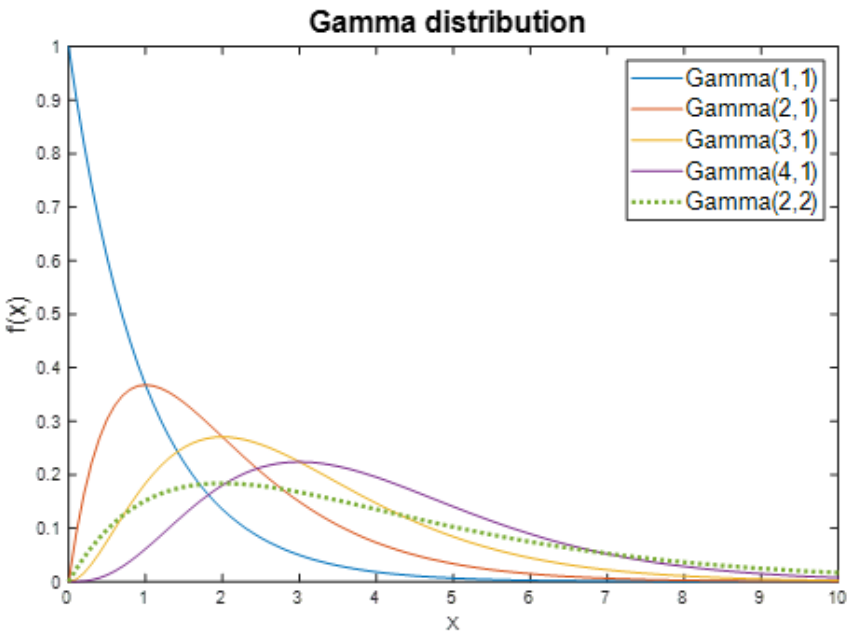
$$f(x) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}, & x > 0 \\ 0, & \textit{otherwise} \end{cases}$$

We denote this by Gamma(α, β) or $\Gamma(\alpha, \beta)$.
*The parameter α is called a **shape** parameter, and β is called a **scale** parameter.*

GAMMA DISTRIBUTION: CHARACTERISTICS

If $X \sim \Gamma(\alpha, \beta)$:

$$f(x) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}, & x > 0 \\ 0, & \text{otherwise} \end{cases}$$



GAMMA DISTRIBUTION: CHARACTERISTICS

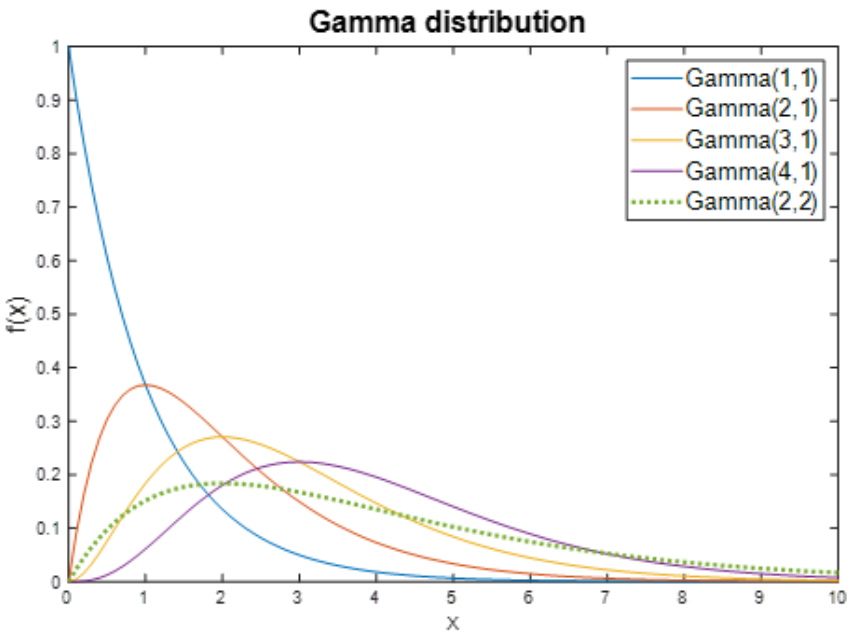
If $X \sim \Gamma(\alpha, \beta)$:

$$f(x) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}, & x > 0 \\ 0, & \text{otherwise} \end{cases}$$

$$E(X) = \alpha\beta, \qquad Var(X) = \alpha\beta^2$$

and the moment generating function:

$$M_X(t) = (1 - \beta t)^{-\alpha}$$



GAMMA DISTRIBUTION: CHARACTERISTICS

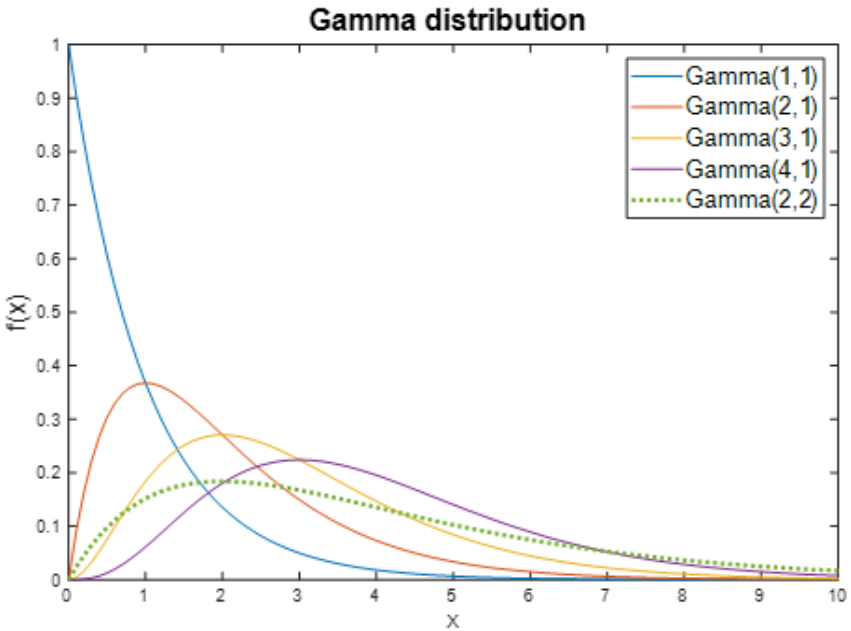
If $X \sim \Gamma(\alpha, \beta)$:

$$f(x) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}, & x > 0 \\ 0, & \text{otherwise} \end{cases}$$

$$E(X) = \alpha\beta, \qquad Var(X) = \alpha\beta^2$$

and the moment generating function:

$$M_X(t) = (1 - \beta t)^{-\alpha}$$



A special case of the gamma distribution is the exponential distribution, if $\alpha = 1, \beta = 1/\lambda$: $\qquad f(x) = \lambda e^{-\lambda x}$

$f(x) = \frac{x}{\beta} e^{-\frac{x}{\beta}}$

GAMMA DISTRIBUTION: CHARACTERISTICS

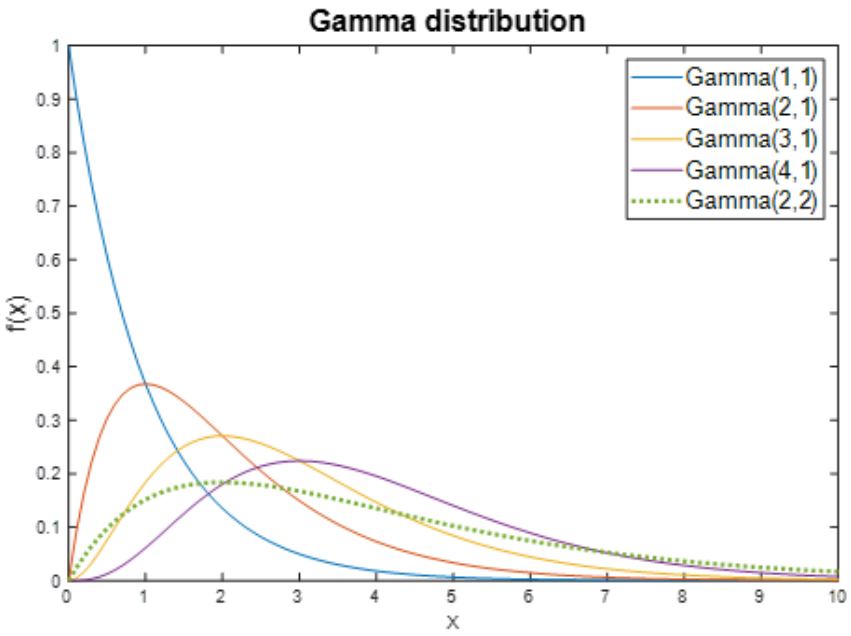
If $X \sim \Gamma(\alpha, \beta)$:

$$f(x) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}, & x > 0 \\ 0, & \text{otherwise} \end{cases}$$

$$E(X) = \alpha\beta, \qquad Var(X) = \alpha\beta^2$$

and the moment generating function:

$$M_X(t) = (1 - \beta t)^{-\alpha}$$



A special case of the gamma distribution is the exponential distribution, if $\alpha = 1, \beta = 1/\lambda$: $\qquad f(x) = \lambda e^{-\lambda x}$

MATLAB: b=1; a=2; x=0:0.1:10; y= gampdf(x,a,b); plot(x,y)
R: a<−4; b<−1; x<−seq(0,10,0.1);
y<−dgamma(x, shape=a, scale = b, log = FALSE); plot(x,y)

χ^2 DISTRIBUTION

DEFINITION

Let $Z \sim N(0,1)$. Then, if $X = Z^2$, we say that X follows the ***chi-square distribution with 1 degree of freedom****.*
We write, $X \sim \chi^2_1$.

Derivation of the χ^2 pdf: `https://en.wikipedia.org/wiki/Proofs_related_to_chi-squared_distribution`

χ^2 DISTRIBUTION

DEFINITION

Let $Z \sim N(0,1)$. Then, if $X = Z^2$, we say that X follows the ***chi-square distribution with 1 degree of freedom****.*
We write, $X \sim \chi^2_1$.

DEFINITION

Let $Z_1, Z_2, \ldots, Z_n$ be independent random variables with $Z_i \sim N(0,1)$. If $Y = \sum_{i=1}^{n} Z_i^2$, then Y follows the ***chi-square distribution with n degrees of freedom****.*
We write, $Y \sim \chi^2_n$.

Derivation of the χ^2 pdf: `https://en.wikipedia.org/wiki/Proofs_related_to_chi-squared_distribution`

χ^2 DISTRIBUTION: CHARACTERISTICS

$\chi_n^2 \sim \Gamma(\frac{n}{2}, 2)$

$n \to \infty$

$\chi_n^2 \to N(n, \sqrt{2n})$

If $X \sim \chi_n^2$, then the probability density distribution is:

$$f(x) = \begin{cases} \frac{1}{2^{n/2}\Gamma(n/2)} x^{n/2-1} e^{-x/2}, & x > 0 \\ 0, \text{ otherwise} \end{cases}$$

Expectation and Variance

$$E(X) = n, \quad Var(X) = 2n,$$

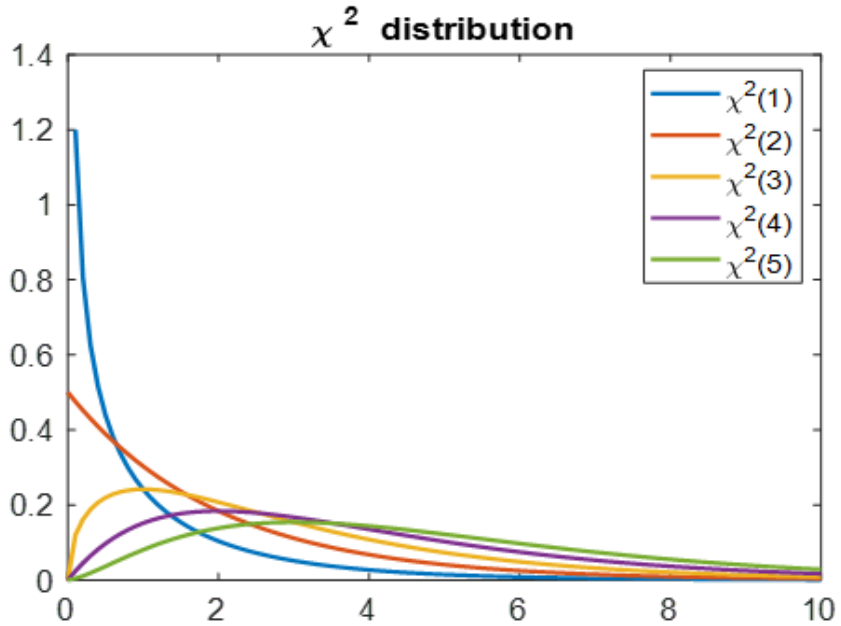
$\alpha\beta$ $\quad$ $\alpha\beta^2$

k-th moment:

$$E(X^k) = \frac{2^k \Gamma[(n/2) + k]}{\Gamma(n/2)},$$

Moment Generating Function

$$M(t) = (1 - 2t)^{-n/2} \text{ for } t < \frac{1}{2}.$$



χ^2 DISTRIBUTION: CHARACTERISTICS

If $X \sim \chi^2_n$, then the probability density distribution is:

$$f(x) = \begin{cases} \frac{1}{2^{n/2}\Gamma(n/2)} x^{n/2-1} e^{-x/2}, & x > 0 \\ 0, \text{otherwise} \end{cases}$$

Expectation and Variance

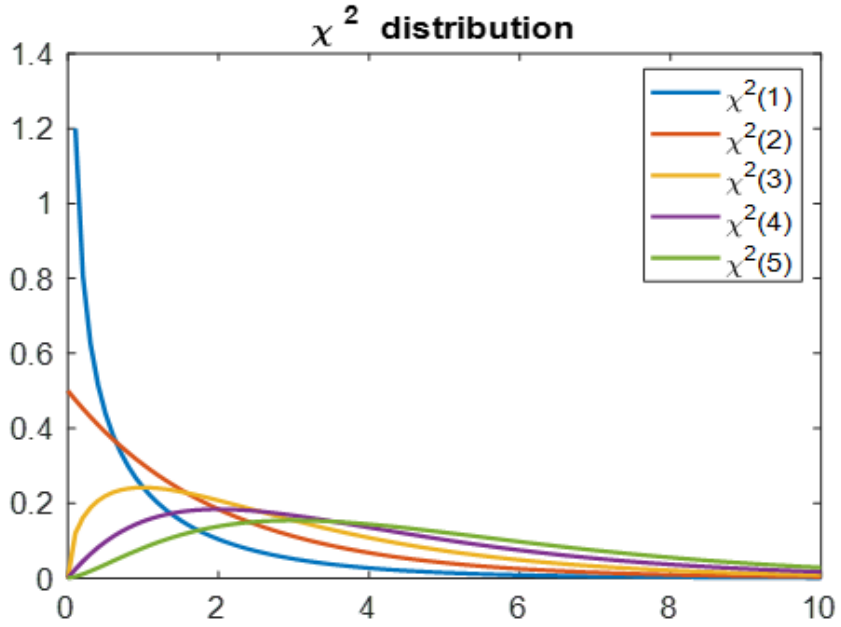
$$E(X) = n, \quad Var(X) = 2n,$$

k-th moment:

$$E(X^k) = \frac{2^k \Gamma[(n/2)+k]}{\Gamma(n/2)},$$

Moment Generating Function

$$M(t) = (1-2t)^{-n/2} \text{ for } t < \frac{1}{2}.$$



MATLAB: n=1, x=0:0.1:10; y= chi2pdf(x,n); plot(x,y)
R: n<−3; x<−seq(0,10,0.1); y<−dchisq(x, fd=n,); plot(x,y)

χ^2 DISTRIBUTION:PROPERTIES

Proposition

Let $X_1, X_2, ...X_n$ be independent normally distributed random variables such that $X_i \sim N(\mu_i, \sigma_i^2)$. Let $Y = c_1X_1 + c_2X_2 + ... + c_nX_n$. Then $Y \sim N(\sum_{i=1}^n c_i\mu_i, \sum_{i=1}^n c_i^2\sigma_i^2)$.

χ^2 DISTRIBUTION:PROPERTIES

THEOREM

Let $X_1, X_2, ..., X_n$ be independent RVs such that $X_j \sim \Gamma(\alpha_j, \beta)$, $j = 1, 2, ..., n$. Then $Y = \sum_{k=1}^{n} X_k$ is a $\Gamma(\sum_{j=1}^{n} \alpha_j, \beta)$ RV.

χ^2 DISTRIBUTION:PROPERTIES

THEOREM

Let $X_1, X_2, ..., X_n$ be independent RVs such that $X_j \sim \Gamma(\alpha_j, \beta)$, $j = 1, 2, ..., n$. Then $Y = \sum_{k=1}^{n} X_k$ is a $\Gamma(\sum_{j=1}^{n} \alpha_j, \beta)$ RV.

COROLLARY (1)

Let $X_1, X_2, ..., X_n$ be iid RVs, each with an exponential distribution with parameter β. Then Y is a $\Gamma(n, \beta)$ RV.

$\Gamma(1, \beta)$

χ^2 DISTRIBUTION:PROPERTIES

THEOREM

Let $X_1, X_2, ..., X_n$ be independent RVs such that $X_j \sim \Gamma(\alpha_j, \beta)$, $j = 1, 2, ..., n$. Then $Y = \sum_{k=1}^{n} X_k$ is a $\Gamma(\sum_{j=1}^{n} \alpha_j, \beta)$ RV.

COROLLARY (1)

Let $X_1, X_2, ..., X_n$ be iid RVs, each with an exponential distribution with parameter β. Then Y is a $\Gamma(n, \beta)$ RV.

COROLLARY (2)

If $X_1, X_2, ..., X_n$ are iid RVs such that $X_j \sim \chi^2(r_j)$, $j = 1, 2, ..., n$, then Y is a $\chi^2_{(\sum_{i=1}^{n} r_j)}$ RV.

$X_j = \sum_{i=1}^{r_j} Z_i^2$

χ^2 DISTRIBUTION:PROPERTIES

THEOREM

Let $X_1, X_2, ..., X_n$ be independent RVs such that $X_j \sim \Gamma(\alpha_j, \beta)$, $j = 1, 2, ..., n$. Then $Y = \sum_{k=1}^{n} X_k$ is a $\Gamma(\sum_{j=1}^{n} \alpha_j, \beta)$ RV.

COROLLARY (1)

Let $X_1, X_2, ..., X_n$ be iid RVs, each with an exponential distribution with parameter β. Then Y is a $\Gamma(n, \beta)$ RV.

COROLLARY (2)

If $X_1, X_2, ..., X_n$ are iid RVs such that $X_j \sim \chi^2(r_j)$, $j = 1, 2, ..., n$, then Y is a $\chi^2_{(\sum_{i=1}^{n} r_j)}$ RV.

COROLLARY (3)

If X_1, X_2 are iid RVs such that $X_j \sim \chi^2(r_j)$, $j = 1, 2$ and $r_1 > r_2$, then $Y = \underline{X_1 - X_2}$ is a $\chi^2_{(r_1 - r_2)}$ RV.

χ^2 DISTRIBUTION:PROPERTIES

$$X \sim N(\mu, \sigma^2) \Rightarrow Z = \frac{X-\mu}{\sigma} \sim N(0,1)$$

THEOREM

Let $X_1, X_2, ..., X_n$ independent random variables with $X_i \sim N(\mu; \sigma^2)$. It follows directly form the previous theorem that if $Y = \sum_{i=1}^{n}(X_i - \mu)^2/\sigma^2$, then $Y \sim \chi^2_n$

χ^2 DISTRIBUTION:PROPERTIES

THEOREM

Let $X_1, X_2, ..., X_n$ independent random variables with $X_i \sim N(\mu; \sigma^2)$. It follows directly form the previous theorem that if $Y = \sum_{i=1}^{n}(X_i - \mu)^2/\sigma^2$, then $Y \sim \chi^2_n$

THEOREM

Let $X_1, X_2, ..., X_n$ independent random variables with $X_i \sim N(\mu; \sigma^2)$. Define the sample variance as $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$. Then

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

χ^2 DISTRIBUTION:PROPERTIES

PROOF.

$$\left[\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}\right]$$

χ^2 DISTRIBUTION:PROPERTIES

PROOF.

$$\frac{(n-1)}{\sigma^2}S^2 = \frac{\cancel{(n-1)}}{\sigma^2}\cdot\frac{1}{\cancel{n-1}}\sum_{i=1}^{n}(X_i - \overset{-\mu+\mu}{\bar{X}})^2 = \frac{1}{\sigma^2}\sum_{i=1}^{n}((X_i-\mu)-(\bar{X}-\mu))^2$$

$$= \sum_{i=1}^{n}\frac{(X_i-\mu)^2}{\sigma^2} - 2\sum_{i=1}^{n}\frac{(X_i-\mu)}{\sigma}\frac{(\bar{X}-\mu)}{\sigma} + \sum_{i=1}^{n}\frac{(\bar{X}-\mu)^2}{\sigma^2}$$

χ^2 DISTRIBUTION:PROPERTIES

PROOF.

$$\frac{(n-1)}{\sigma^2}S^2 = \frac{(n-1)}{\sigma^2}\cdot\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2 = \frac{1}{\sigma^2}\sum_{i=1}^{n}((X_i-\mu)-(\bar{X}-\mu))^2$$

$$= \sum_{i=1}^{n}\frac{(X_i-\mu)^2}{\sigma^2} - 2\sum_{i=1}^{n}\frac{(X_i-\mu)}{\sigma}\frac{(\bar{X}-\mu)}{\sigma} + \sum_{i=1}^{n}\frac{(\bar{X}-\mu)^2}{\sigma^2}$$

Denote: $W_1 = \sum_{i=1}^{n}\frac{(X_i-\mu)^2}{\sigma^2}$, $W_2 = \sum_{i=1}^{n}\frac{(X_i-\mu)}{\sigma}\frac{(\bar{X}-\mu)}{\sigma}$, $W_3 = \sum_{i=1}^{n}\frac{(\bar{X}-\mu)^2}{\sigma^2}$.

χ^2 DISTRIBUTION:PROPERTIES

PROOF.

$$\frac{(n-1)}{\sigma^2}S^2 = \frac{(n-1)}{\sigma^2}\cdot\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2 = \frac{1}{\sigma^2}\sum_{i=1}^{n}((X_i-\mu)-(\bar{X}-\mu))^2$$

$$= \sum_{i=1}^{n}\frac{(X_i-\mu)^2}{\sigma^2} - 2\sum_{i=1}^{n}\frac{(X_i-\mu)}{\sigma}\frac{(\bar{X}-\mu)}{\sigma} + \sum_{i=1}^{n}\frac{(\bar{X}-\mu)^2}{\sigma^2}$$

Denote: $W_1 = \sum_{i=1}^{n}\frac{(X_i-\mu)^2}{\sigma^2}$, $W_2 = \sum_{i=1}^{n}\frac{(X_i-\mu)}{\sigma}\frac{(\bar{X}-\mu)}{\sigma}$, $W_3 = \sum_{i=1}^{n}\frac{(\bar{X}-\mu)^2}{\sigma^2}$.

W_1: $\frac{(X_i-\mu)}{\sigma} \sim N(0,1)$, hence, $W_1 \sim \chi^2_n$.

χ^2 DISTRIBUTION:PROPERTIES

PROOF.

$$\frac{(n-1)}{\sigma^2}S^2 = \frac{(n-1)}{\sigma^2}\cdot\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2 = \frac{1}{\sigma^2}\sum_{i=1}^{n}((X_i-\mu)-(\bar{X}-\mu))^2$$

$$= \sum_{i=1}^{n}\frac{(X_i-\mu)^2}{\sigma^2} - 2\sum_{i=1}^{n}\frac{(X_i-\mu)}{\sigma}\frac{(\bar{X}-\mu)}{\sigma} + \sum_{i=1}^{n}\frac{(\bar{X}-\mu)^2}{\sigma^2}$$

Denote: $W_1 = \sum_{i=1}^{n}\frac{(X_i-\mu)^2}{\sigma^2}$, $W_2 = \sum_{i=1}^{n}\frac{(X_i-\mu)}{\sigma}\frac{(\bar{X}-\mu)}{\sigma}$, $W_3 = \sum_{i=1}^{n}\frac{(\bar{X}-\mu)^2}{\sigma^2}$.

W_1: $\frac{(X_i-\mu)}{\sigma} \sim N(0,1)$, hence, $W_1 \sim \chi^2_n$.

W_2, W_3: $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$, hence $\frac{(\bar{X}-\mu)}{\sigma/\sqrt{n}} \sim N(0,1) \Rightarrow$ denote $Y = \frac{(\bar{X}-\mu)^2}{\sigma^2/n} \sim \chi^2_1$

$W_3 = \sum_{i=1}^{n}\frac{(\bar{X}-\mu)^2}{\sigma^2} = n\frac{(\bar{X}-\mu)^2}{\sigma^2} = \frac{(\bar{X}-\mu)^2}{\sigma^2/n} = Y.$

W_2: $\sum_{i=1}^{n}\frac{(X_i-\mu)}{\sigma}\frac{(\bar{X}-\mu)}{\sigma} = \frac{(\bar{X}-\mu)}{\sigma}\cdot\frac{n}{n}\cdot\sum_{i=1}^{n}\frac{(X_i-\mu)}{\sigma} = \frac{(\bar{X}-\mu)}{\sigma}n\frac{(\bar{X}-\mu)}{\sigma} = n\frac{(\bar{X}-\mu)^2}{\sigma^2} = Y$

χ^2 DISTRIBUTION:PROPERTIES

PROOF.

$$\frac{(n-1)}{\sigma^2}S^2 = \frac{(n-1)}{\sigma^2}\cdot\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2 = \frac{1}{\sigma^2}\sum_{i=1}^{n}((X_i-\mu)-(\bar{X}-\mu))^2$$

$$= \sum_{i=1}^{n}\frac{(X_i-\mu)^2}{\sigma^2} - 2\sum_{i=1}^{n}\frac{(X_i-\mu)}{\sigma}\frac{(\bar{X}-\mu)}{\sigma} + \sum_{i=1}^{n}\frac{(\bar{X}-\mu)^2}{\sigma^2}$$

$-2y + y = -y$

Denote: $W_1 = \sum_{i=1}^{n}\frac{(X_i-\mu)^2}{\sigma^2}$, $W_2 = \sum_{i=1}^{n}\frac{(X_i-\mu)}{\sigma}\frac{(\bar{X}-\mu)}{\sigma}$, $W_3 = \sum_{i=1}^{n}\frac{(\bar{X}-\mu)^2}{\sigma^2}$.

W_1: $\frac{(X_i-\mu)}{\sigma} \sim N(0,1)$, hence, $W_1 \sim \chi^2_n$.

W_2, W_3: $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$, hence $\frac{(\bar{X}-\mu)}{\sigma/\sqrt{n}} \sim N(0,1) \Rightarrow$ denote $Y = \frac{(\bar{X}-\mu)^2}{\sigma^2/n} \sim \chi^2_1$

$$W_3 = \sum_{i=1}^{n}\frac{(\bar{X}-\mu)^2}{\sigma^2} = n\frac{(\bar{X}-\mu)^2}{\sigma^2} = \frac{(\bar{X}-\mu)^2}{\sigma^2/n} = Y.$$

$$W_2: \sum_{i=1}^{n}\frac{(X_i-\mu)}{\sigma}\frac{(\bar{X}-\mu)}{\sigma} = \frac{(\bar{X}-\mu)}{\sigma}\frac{n}{n}\sum_{i=1}^{n}\frac{(X_i-\mu)}{\sigma} = \frac{(\bar{X}-\mu)}{\sigma}n\frac{(\bar{X}-\mu)}{\sigma} = n\frac{(\bar{X}-\mu)^2}{\sigma^2} = Y$$

$$\frac{(n-1)S^2}{\sigma^2} = W_1 - Y \sim \chi^2_{n-1}$$

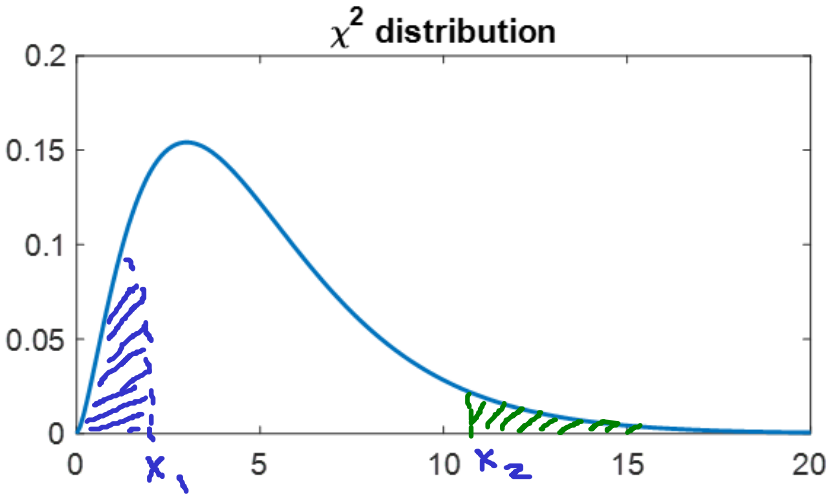
□

χ^2 DISTRIBUTION: INTERVAL ESTIMATION

If $X \sim \chi^2_n$, then from the χ^2 table, we can compute the values of $\chi^2_{\alpha,\,n}$ (or $\chi^2_\alpha(n)$) such that

$$P(X > \chi^2_{\alpha,\,n}) = \alpha \qquad \Leftrightarrow \qquad P(X \le \chi^2_{\alpha,\,n}) = 1 - \alpha$$

	$P(X \le x)$							
	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990
r	$\chi^2_{0.99}(r)$	$\chi^2_{0.975}(r)$	$\chi^2_{0.95}(r)$	$\chi^2_{0.90}(r)$	$\chi^2_{0.10}(r)$	$\chi^2_{0.05}(r)$	$\chi^2_{0.025}(r)$	$\chi^2_{0.01}(r)$
1	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635
2	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210
3	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.34
4	0.297	0.484	0.711	1.064	7.779	9.488	11.14	13.28
5	0.554	0.831	1.145	1.610	9.236	11.07	12.83	15.09
6	0.872	1.237	1.635	2.204	10.64	12.59	14.45	16.81
7	1.239	1.690	2.167	2.833	12.02	14.07	16.01	18.48
8	1.646	2.180	2.733	3.490	13.36	15.51	17.54	20.09
9	2.088	2.700	3.325	4.168	14.68	16.92	19.02	21.67
10	2.558	3.247	3.940	4.865	15.99	18.31	20.48	23.21



$X \sim \chi^2_6$

$P(X < x_1) = 0{,}1$

$P(X < \chi^2_{0{,}9,6}) = P(X < 2{,}2) = 0{,}1$

$P(X > x_2) = 0{,}1$

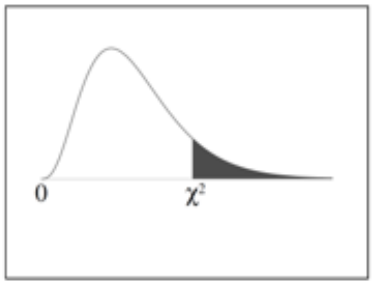
$P(X > \chi^2_{0{,}1,6}) = P(X > 10{,}64) = 0{,}1$

χ^2 DISTRIBUTION: INTERVAL ESTIMATION

If $X \sim \chi^2_n$, then from the χ^2 table, we can compute the values of $\chi^2_{\alpha,\, n}$ (or $\chi^2_\alpha(n)$) such that

$$P(X > \chi^2_{\alpha,\, n}) = \alpha \qquad \Leftrightarrow \qquad P(X \leq \chi^2_{\alpha,\, n}) = 1 - \alpha$$

Chi-Square Distribution Table



The shaded area is equal to α for $\chi^2 = \chi^2_\alpha$.

df	$\chi^2_{.995}$	$\chi^2_{.990}$	$\chi^2_{.975}$	$\chi^2_{.950}$	$\chi^2_{.900}$	$\chi^2_{.100}$	$\chi^2_{.050}$	$\chi^2_{.025}$	$\chi^2_{.010}$	$\chi^2_{.005}$
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188

χ^2 DISTRIBUTION: INTERVAL ESTIMATION

If $X \sim \chi^2_n$, then from the χ^2 table, we can compute the values of $\chi^2_{\alpha,\, n}$ (or $\chi^2_\alpha(n)$) such that

$$P(X > \chi^2_{\alpha,\, n}) = \alpha \qquad \Leftrightarrow \qquad P(X \leq \chi^2_{\alpha,\, n}) = 1 - \alpha$$

EXAMPLE

Let the random variables $X_1, X_2, \ldots, X_5$ be from an $N(5, 1)$ distribution. Find a number ε such that $P_X\left(\sum_{i=1}^5 (X_i - 5)^2 \leq \varepsilon\right) = 0.90$?

$\frac{X_i - 5}{1} \sim N(0,1)$

Solution:

$\sum_{i=1}^5 (X_i - 5)^2 \sim \chi^2_{0.1,\, 5}$ looking in the table with 5 df for $\chi^2_{0.10}$ we obtain, $\chi^2_{0.1,5}$

$\varepsilon = 9.23635.$

Alternatively, we can find the value in MATLAB: eps=chi2inv(0.9, 5)
or in R: eps<−qchisq(0.9, df=5)

χ^2 DISTRIBUTION: INTERVAL ESTIMATION

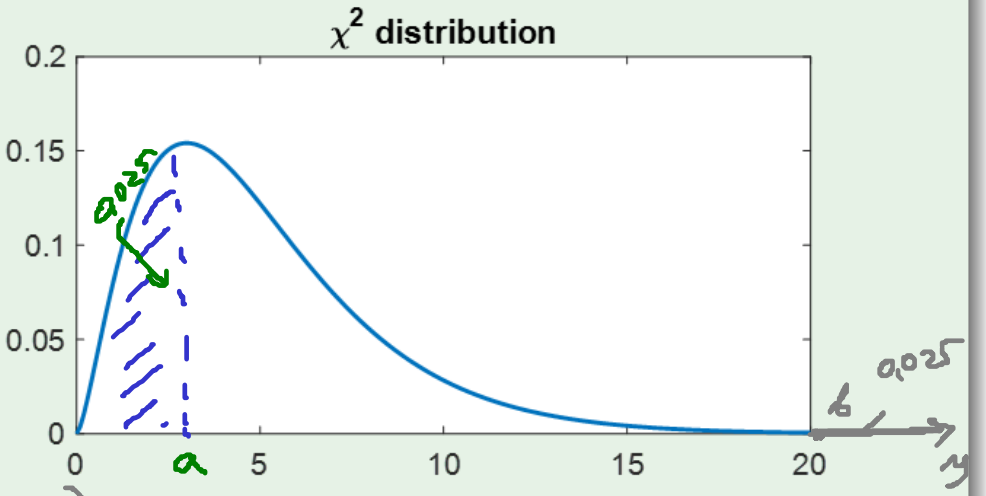
EXAMPLE

Let $X_1, X_2, \ldots, X_{10}$ be a random sample from a standard normal distribution. Find the numbers a and b such that

$$P(a \leq \underbrace{\sum_{i=1}^{10} X_i^2}_{y \sim \chi^2_{10}} \leq b) = 0.95$$

Solution:

	$P(X \leq x)$							
	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990
r	$\chi^2_{0.99}(r)$	$\chi^2_{0.975}(r)$	$\chi^2_{0.95}(r)$	$\chi^2_{0.90}(r)$	$\chi^2_{0.10}(r)$	$\chi^2_{0.05}(r)$	$\chi^2_{0.025}(r)$	$\chi^2_{0.01}(r)$
1	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635
2	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210
3	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.34
4	0.297	0.484	0.711	1.064	7.779	9.488	11.14	13.28
5	0.554	0.831	1.145	1.610	9.236	11.07	12.83	15.09
6	0.872	1.237	1.635	2.204	10.64	12.59	14.45	16.81
7	1.239	1.690	2.167	2.833	12.02	14.07	16.01	18.48
8	1.646	2.180	2.733	3.490	13.36	15.51	17.54	20.09
9	2.088	2.700	3.325	4.168	14.68	16.92	19.02	21.67
10	2.558	3.247	3.940	4.865	15.99	18.31	20.48	23.21



$P(y > b) = 0{,}025 = P(y > \chi^2_{0{,}025,10}) = P(y > 20{,}48)$

$P(y < a) = 0{,}025 = P(y < \chi^2_{0{,}975,10}) = P(y < 3{,}247)$

$P(3{,}247 < \sum_{i=1}^{10} X_i^2 < 20{,}48) = 0{,}95$

χ^2 DISTRIBUTION: INTERVAL ESTIMATION

Procedure to find confidence interval for σ^2:

1. Calculate $\bar{x}$ and s^2 from the sample $x_1, \ldots, x_n$.

χ^2 DISTRIBUTION: INTERVAL ESTIMATION

Procedure to find confidence interval for σ^2:

1. Calculate $\bar{x}$ and s^2 from the sample $x_1, ..., x_n$.
2. Find $L = \chi^2_{\alpha/2}$, and $U = \chi^2_{1-\alpha/2}$ using the χ^2 table with $(n-1)$ degrees of freedom.

χ^2 DISTRIBUTION: INTERVAL ESTIMATION

Procedure to find confidence interval for σ^2:

1. Calculate $\bar{x}$ and s^2 from the sample $x_1, \ldots, x_n$.
2. Find $L = \chi^2_{\alpha/2}$, and $U = \chi^2_{1-\alpha/2}$ using the χ^2 table with $(n-1)$ degrees of freedom.
3. Compute the $(1-\alpha)100\%$ confidence interval for the population variance s^2 as
$$\left(\frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}}, \frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}}\right),$$
where χ^2-values are with $(n-1)$ degrees of freedom.

χ^2 DISTRIBUTION: INTERVAL ESTIMATION

Procedure to find confidence interval for σ^2:

1. Calculate $\bar{x}$ and s^2 from the sample $x_1, \ldots, x_n$.
2. Find $L = \chi^2_{\alpha/2}$, and $U = \chi^2_{1-\alpha/2}$ using the χ^2 table with $(n-1)$ degrees of freedom.
3. Compute the $(1-\alpha)100\%$ confidence interval for the population variance s^2 as
$$\left(\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}, \frac{(n-1)s^2}{\chi^2_{\alpha/2}}\right),$$
where χ^2-values are with $(n-1)$ degrees of freedom.

Assumption: The population is normal.

χ^2 DISTRIBUTION: INTERVAL ESTIMATION

EXAMPLE

Suppose we have an independent random sample $X_1, ..., X_{10}$ from $N(\mu; \sigma^2)$ and we want to find a 95% confidence interval for σ^2.

Solution:

1) As $(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$, the pivot $T = 9S^2/\sigma^2 \sim \chi^2_9$.

2) Using χ^2 table find the lower bound $L = \chi^2_{\alpha/2,n-1}$ and the upper bound $U = \chi^2_{1-\alpha/2,n-1}$: $L = 2.70$, $U = 19.02$, then

$$P(L < T < U) = P\left(L < \frac{(n-1)S^2}{\sigma^2} < U\right) = 1 - \alpha$$

$$P\left(\frac{(n-1)S^2}{U} < \sigma^2 < \frac{(n-1)S^2}{L}\right) = 1 - \alpha$$

$$P\left(\frac{9S^2}{19.02} < \sigma^2 < \frac{9S^2}{2.70}\right) = 0.95$$

The confidence interval $$\left(\frac{9S^2}{19.02}, \frac{9S^2}{2.70}\right)$$

χ^2 DISTRIBUTION: INTERVAL ESTIMATION

EXAMPLE

A fruit-drink company wants to know the variation, as measured by the standard deviation, of the amount of juice in cans. From past experience, it is known that $\sigma^2 = 2$. The company statistician decides to take a sample of 25 cans from the production line and compute the sample variance.
Assuming that the sample values may be viewed as a random sample from a normal population, find a value of b such that $P(S^2 > b) = 0.05$.

χ^2 DISTRIBUTION: INTERVAL ESTIMATION

EXAMPLE

A fruit-drink company wants to know the variation, as measured by the standard deviation, of the amount of juice in cans. From past experience, it is known that $\sigma^2 = 2$. The company statistician decides to take a sample of 25 cans from the production line and compute the sample variance.
Assuming that the sample values may be viewed as a random sample from a normal population, find a value of b such that $P(S^2 > b) = 0.05$.
Solution:
$(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$, hence,

$\sigma^2 = 2 \neq 0$

$n-1 = 24$

$$0.05 = P(\underbrace{\frac{24S^2}{\sigma^2}}_{X} > \frac{24b}{\sigma^2}) = P(X > c), \quad X \sim \chi^2_{24}$$

We obtain $c = 36.415$, thus $b = c \cdot 2/24 = 3.035$,

$$P(S^2 > 3.035) = 0.05$$

χ^2 DISTRIBUTION: INTERVAL ESTIMATION

EXAMPLE

A fruit-drink company wants to know the variation, as measured by the standard deviation, of the amount of juice in cans. σ^2 is unknown.
The company statistician decided to take a sample of 28 cans from the production line and observed that the sample variance is $s^2 = 4.5$.
Assume that the sample values may be viewed as a random sample from a normal population. At which level of confidence we can say that $\sigma^2 \leq 5$.

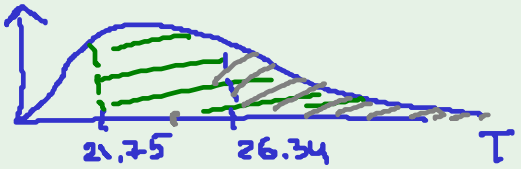
χ^2 DISTRIBUTION: INTERVAL ESTIMATION

EXAMPLE

A fruit-drink company wants to know the variation, as measured by the standard deviation, of the amount of juice in cans. σ^2 is unknown.
The company statistician decided to take a sample of 28 cans from the production line and observed that the sample variance is $s^2 = 4.5$.
Assume that the sample values may be viewed as a random sample from a normal population. At which level of confidence we can say that $\sigma^2 \leq 5$.

Solution:

$(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$, hence,



$$P(\sigma^2 \leq 5) = P(\frac{27s^2}{5} \leq \frac{27s^2}{\sigma^2}) = P(\frac{27s^2}{5} \leq X), \quad X \sim \chi^2_{27}$$

As $\frac{27s^2}{5} = 24.3$, from the table $P(X \geq 21.75) > P(X \geq 24.3) > P(X \geq 26.34)$,

$$0.5 < P(X \geq 24.3) < 0.75 \qquad \Leftrightarrow \qquad 0.5 < P(\sigma^2 \leq 5) < 0.75$$

We can obtain exact value using a software, for example, in R:
alpha<−pchisq(24.3, df=27), hence, $1 - \alpha = 0.613$, i.e. 61%.

APPENDIX

Very interesting discussion about relation between some distributions, in particular Normal, χ^2 and Gamma.
Please look at the answer provided by Carl

https://stats.stackexchange.com/questions/37461/the-relationship-between-the-gamma-distribution-and-the-normal-di

QUESTIONS TO TAKE HOME

1. Define χ_1^2, χ_n^2 distributions and their characteristics;

QUESTIONS TO TAKE HOME

1. Define χ_1^2, χ_n^2 distributions and their characteristics;
2. State the theorems about the sum and the difference of random variables from χ^2 distributions (Corollary 2, 3);

QUESTIONS TO TAKE HOME

1. Define χ_1^2, χ_n^2 distributions and their characteristics;
2. State the theorems about the sum and the difference of random variables from χ^2 distributions (Corollary 2, 3);
3. Define the sample variance;

QUESTIONS TO TAKE HOME

1. Define χ_1^2, χ_n^2 distributions and their characteristics;
2. State the theorems about the sum and the difference of random variables from χ^2 distributions (Corollary 2, 3);
3. Define the sample variance;
4. State and proof the theorem about the distribution of the sample variance;

QUESTIONS TO TAKE HOME

1. Define χ_1^2, χ_n^2 distributions and their characteristics;
2. State the theorems about the sum and the difference of random variables from χ^2 distributions (Corollary 2, 3);
3. Define the sample variance;
4. State and proof the theorem about the distribution of the sample variance;
5. Define the confidence interval for the population variance.