MATH 321: Test 2 Study Guide

Lecture 1 – Random Samples and Common Statistics (5.5)

Basic concepts of random samples

- Random sample definition: $X_1, \ldots, X_n$ are a random sample of size n from the population $f(x)$ if they are *iid* random variables.
- Statistic (estimator) definition: The random variable / vector for any function of a random sample $Y = T(X_1, \ldots, X_n)$ is called a statistic, and it's distribution is called a sampling distribution.

Sample mean and variance

- Definitions
  - Sample mean: The arithmetic average of the values in a random sample

    $$\bar{X} = \frac{X_1 + \cdots + X_n}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

  - Sample variance: The statistic defined by $S^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}(X_i - \bar{X})^2$
  - Sample standard deviation: The statistic defined by $S = \sqrt{S^2}$
- Theorem: Let $X_1, \ldots, X_n$ be a random sample of size n from a population with mean μ and variance $\sigma^2 < \infty$. Then

  (a) $\mu_{\bar{X}} = E(\bar{X}) = \mu$ (b) $\sigma^2_{\bar{X}} = V(\bar{X}) = \frac{\sigma^2}{n}$ (c) $E(S^2) = \sigma^2$
- Sampling distribution of $\bar{X}$ from random sample $X_1, \ldots, X_n$

  Theorem: Mgf of the sample mean is $M_{\bar{X}}(t) = [M_X(t/n)]^n$

Sampling from the normal distribution

- Let $X_1, \ldots, X_n$ be a random sample of size n from a Normal (μ, σ^2) distribution. Then

  (a) $\bar{X} \perp\!\!\!\perp S^2$ (b) $\bar{X} \sim \text{Normal}\,(\mu, \frac{\sigma^2}{n})$ (c) $\dfrac{(n-1)}{\sigma^2}S^2 \sim \chi^2\,(n-1)$

Chi-square random variables

- If $Z \sim \text{Normal}\,(0,1)$, then $Z^2 \sim \chi^2\,(1) \rightarrow \left(\dfrac{\bar{X}-\mu}{\sigma}\right)^2 = Z^2 \sim \chi^2\,(1)$
- Additive *df*: If $X_1, \ldots, X_n$ are mutually independent and $X_i \sim \chi^2\,(r_i)$ for $i = 1, \ldots, n$, then $Y = X_1 + \cdots + X_n \sim \chi^2\,(r_1 + \cdots + r_n)$
- Result / extension of this: If $X_1, \ldots, X_n$ are mutually independent random variables with $X_i \sim \text{Normal}\,(\mu_i, \sigma_i)$ for $i = 1, \ldots, n$, then

  $$\sum_{i=1}^{n}\left(\frac{\bar{X}-\mu}{\sigma}\right)^2 = \sum_{i=1}^{n} Z^2 \sim \chi^2\,(n)$$

t distribution

- Definition: Let $X_1, \dots, X_n$ be a random sample from a $N(\mu, \sigma^2)$ population. Then $\dfrac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{\,n-1}$
- Derivation: $\dfrac{Z}{\sqrt{{\chi^2}_r / r}} \sim t_{\,r}$

F distribution

- Definition: Let $X_1, \dots, X_n$ be a random sample from a $N(\mu_X, \sigma_X^2)$ population, and let $Y_1, \dots, Y_m$ be a random sample from an independent $N(\mu_Y, \sigma_Y^2)$ population. If

  $W = \dfrac{S_X^2 / S_Y^2}{\sigma_X^2 / \sigma_Y^2}$ then $W \sim F\,(n-1, m-1)$. In general, $W \sim F\,(r_1, r_2)$.
- Derivation: $\dfrac{{\chi^2}_{r_1} / r_1}{{\chi^2}_{r_2} / r_2} \sim F\,(r_1, r_2)$
- Relationship to other distributions theorem

  (a) If $X \sim F\,(r_1, r_2)$ then $1/X \sim F\,(r_2, r_1)$ (b) If $X \sim t_{\,r}$ then $X^2 \sim F\,(1, r)$

Lecture 2 – Order Statistics (6.3)

Order statistics definition and distributions

- Definition: The order statistics are random variables that satisfy $X_{(1)} \le \cdots \le X_{(n)}$. In particular

$$
\begin{aligned}
X_{(1)} &= \min_{1 \le i \le n} X_i, \\
X_{(2)} &= \text{second smallest } X_i \\
&\ \ \vdots \\
X_{(n)} &= \max_{1 \le i \le n} X_i.
\end{aligned}
$$

- Distribution theorems
  - Cdf:

$$
\begin{aligned}
F_{X_{(j)}}(x) = P(X_{(j)} \le x) &= \sum_{k=j}^{n} \binom{n}{k} [F_X(x)]^k [1 - F_X(x)]^{n-k} \\
&= P(Y \le j), \quad \text{where} \quad Y \sim \text{Binomial}\,(n, p = P(X \le x) = F_X(x))
\end{aligned}
$$

  - Pdf:

$$
\begin{aligned}
f_{X_{(j)}}(x) &= \frac{n!}{(j-1)!(n-j)!} \, [F_X(x)]^{j-1} \, f_X(x) \, [1 - F_X(x)]^{n-j} \\
&= [\text{multinomial coefficient}] \times [j-1 \text{ RVs } \le x] \times [1 \text{ RV} \approx x] \times [n-j \text{ RVs } > x]
\end{aligned}
$$

- $f_{X_{(j)}}(x) = F'_{X_{(j)}}(x)$

- Extreme order stats

  Min $\rightarrow$ $F_{X_{(1)}}(x) = 1 - [1 - F_X(x)]^n$; $f_{X_{(1)}}(x) = nf_X(x)[1 - F_X(x)]^{n-1}$

  Max $\rightarrow$ $F_{X_{(n)}}(x) = [F_X(x)]^n$; $f_{X_{(n)}}(x) = n[F_X(x)]^{n-1}f_X(x)$

Specific order statistics and functions of order statistics

- Sample median M

$$M = \begin{cases} X_{(\frac{n+1}{2})} & \text{if } n \text{ is odd} \\ \left[X_{(\frac{n}{2})} + X_{(\frac{n}{2}+1)}\right]/2 & \text{if } n \text{ is even} \end{cases}$$

- Sample range, $R = X_{(n)} - X_{(1)} = max(X_1, \ldots, X_n) - min(X_1, \ldots, X_n)$
- $IQR = Q_3 - Q_1$
- Midrange $= \dfrac{X_{(1)} + X_{(n)}}{2}$

Order statistics as estimators of population percentiles

- Expected value of the "position" of order statistics theorem

  If $X_{(1)}, \ldots, X_{(n)}$ are order statistics, then $E[F_X(X_{(j)})] = \dfrac{j}{n+1}, \quad j = 1, \ldots, n$

  Can use $X_{(j)}$ as an estimator of x_p, where $p = j/(n+1)$.

q–q plots

- Expected probability between two adjacent order statistics theorem:

  $E[F_X(X_{(j)}) - F_X(X_{(j-1)})] = \frac{1}{n+1}$; $E[F_X(X_{(1)})] = \frac{1}{n+1}$; $E[1 - F_X(X_{(n)})] = \frac{1}{n+1}$
- q–q plot definition: Let $x_{(1)}, \ldots, x_{(n)}$ be the observed sample order statistics and $x_{\frac{1}{n+1}}, \ldots, x_{\frac{n}{n+1}}$ be the percentiles from some particular distribution. A q–q plot is a plot of the points

  $(x_{(1)}, x_{\frac{1}{n+1}})$, ... , $(x_{(n)}, x_{\frac{n}{n+1}})$
- Interpretation of a q–q plot

  Good model $\rightarrow$ Follows $y = x$ line.

  Bad model $\rightarrow$ Strong deviation from this line.
- q–q plots for the normal distribution.

  If plot $(x_{(1)}, z_{\frac{1}{n+1}})$, ... , $(x_{(n)}, z_{\frac{n}{n+1}})$, then $\frac{1}{\text{slope}} \approx \sigma$

Lecture 3 – Exploratory Data Analysis (6.2)

Univariate EDA

- Descriptive statistics: Goal is to summarize a whole dataset wtih a single or few measures
  - Sample mean $\bar{x} = \dfrac{1}{n}\sum_{i=1}^{n} x_i$
  - Sample variance: $s^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \dfrac{n}{n-1}v$
  - Data (or population) variance: $v = \dfrac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$
- Displaying data
  - Frequency tables: Data is grouped into intervals of equal length (bins)

    Freq = count of observations in each; Relative freq = proportion of observations in each bin = Freq / n
  - Histograms: Shape and summary stats

    $$\begin{array}{rccccc} \text{Right-skewed:} & \text{mean} & > & \text{median} & > & \text{mode} \\ \text{Symmetric:} & \text{mean} & \approx & \text{median} & \approx & \text{mode} \\ \text{Left-skewed:} & \text{mean} & < & \text{median} & < & \text{mode} \end{array}$$
  - Density histograms: Estimate underlying pdf

    For constants c_1 and c_2, $P(c_1 \leq X < c_2) \approx \frac{\text{Freq}}{n}$ on $(c_1, c_2]$

    Height of bar $h(x) = \frac{\text{Freq}}{n(c_2 - c_1)}$
- Empirical rule:
  1. $\approx 68\%$ of data is in $(\bar{x} - s, \bar{x} + s)$.
  2. $\approx 95\%$ of data is in $(\bar{x} - 2s, \bar{x} + 2s)$.
  3. $\approx 99.7\%$ of data is in $(\bar{x} - 3s, \bar{x} + 3s)$.
- Order statistics:
  - 5 number summary
    1. Sample minimum $x_{(1)}$
    2. Lower quartile or First (lower) quartile $q_1 = \hat{x}_{0.25}$
    3. Median (second quartile) $m = \hat{x}_{0.5}$
    4. Third (upper) quartile $q_3 = \hat{x}_{0.75}$
    5. Sample maximum $x_{(n)}$
  - Other statistics

    Sample range, $R = x_{(n)} - x_{(1)}$; $IQR = q_3 - q_1$; Midrange $= \dfrac{x_{(1)} + x_{(n)}}{2}$
  - Boxplots: Visual of 5-number summary, also used to identify outliers

    Suspected outlier → Below $q_1 - 1.5 \times IQR$ (low outlier) or above $q_3 + 1.5 \times IQR$

    Outlier → Below $q_1 - 3 \times IQR$ (low outlier) or above $q_3 + 3 \times IQR$

- Another way to identify outliers: Three-sigma rule

  Outlier if outside $(\bar{x} - 3s, \bar{x} + 3s)$

- q–q plots can be used to test potential models

Bivariate EDA

- Goal: Examine pairwise relationships between variables
- Visualizing dependence: Scatterplots can be used to look for positive, negative or no association.
- Quantifying linear dependence:

  Sample correlation $r = \dfrac{1}{n-1} \dfrac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{s_x\, s_y}$

<u>**Lecture 4 – Point Estimation**</u> (5.8 and 6.4)

Point estimators

- Definition: A point estimator is any function $\hat{\theta} = W(X_1, \ldots, X_n)$ of a sample; that is, any statistic is a point estimator
- An estimator is a random variable (a function of the sample); an estimate is the realized value of the random variable once data is collected

Evaluate estimators

- Unbiased definition: Point estimator $\hat{\theta}$ is unbiased if $E(\hat{\theta}) = \theta$; otherwise it is biased.

  This tells us the mean of a statistic, regardless of n.

- Consistency definition: The property summarized by the WLLN that says if a sequence of the "same" sample quantity approaches a constant as $n \to \infty$, then it is consistent.

  In other words, ff a statistic is consistent, then as $n \to \infty$, there is no variation in what the statistic converges to; the entire distribution converges to a constant.

  - Convergence in probability
    - Definition: A sequence of random variables, $Y_1, Y_2, \ldots,$ converges in probability to a random variable Y if, for every $\epsilon > 0$,

      $$\lim_{n\to\infty} P(|Y_n - Y| \geq \epsilon) = 0 \qquad \text{or, equivalently,} \qquad \lim_{n\to\infty} P(|Y_n - Y| < \epsilon) = 1$$

    - Notation: $Y_n \xrightarrow{p} Y$
  - (Weak) Law of Large Numbers (WLLN)
    - WLLN theorem: Let $X_1, X_2, \ldots$ be *iid* random variable with $E(X_i) = \mu$ and $V(X_i) = \sigma^2 < \infty$. Define $\bar{X}_n = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$. Then for every $\epsilon > 0$,

      $$\lim_{n\to\infty} P(|\bar{X}_n - \mu| < \epsilon) = 1 \qquad\qquad \lim_{n\to\infty} P(|\bar{X}_n - \mu| \geq \epsilon) = 0$$

      that is, $\bar{X} \xrightarrow{p} \mu$.

Method of moments

- Types of moments:
  - k^{th} (population) moment of the distribution (about the origin) $= \mu'_k = E(X^k)$
  - The corresponding sample moment is the average $= m'_k = \dfrac{1}{n}\sum_{i=1}^{n} X_i^k$
- Official statement of Method of Moments:

  Choose as estimates those values of the parameters that are solutions of the equations $\mu'_k = m'_k$, for $k = 1, 2 \ldots, t$, where t is the number of parameters to be estimated
- Steps to find MME
  1. Write $E(X^k)$ as a function of the parameters of interest (may have to integrate)
  2. Then estimate the parameter of interest by equating the population moment with the sample moment and solving for the parameter

Maximum Likelihood Estimation

- Needed items:
  - Parameter space: Set of all possible values for $\theta_1, \ldots, \theta_k$ in pdf (or pmf) $f(x \mid \theta_1, \ldots, \theta_k)$
  - Likelihood function: $L(\boldsymbol{\theta} \mid \mathbf{x}) = f(\mathbf{x} \mid \boldsymbol{\theta}) = \prod_{i=1}^{n} f(x_i \mid \boldsymbol{\theta})$

    Equivalent to the joint pdf or pmf of the data, just with different information considered known.
- MLE definition: For each sample point $\mathbf{x}$, let $\hat{\theta}(\boldsymbol{x})$ be a parameter value at which $L(\theta \mid \mathbf{x})$ attains its maximum as a function of θ, with $\mathbf{x}$ held fixed. A maximum likelihood estimator (MLE) of the parameter θ based on a sample $\mathbf{X}$ is $\hat{\theta}(\mathbf{X})$.
- Steps to find MLEs
  1. Write the likelihood function (i.e. joint density function) and the log-likelihood,

     $$L(\theta \mid \mathbf{x}) = \prod_{i=1}^{n} f(\mathbf{x} \mid \theta) \qquad \rightarrow \qquad \ell(\theta) = \ln[L(\theta \mid \mathbf{x})]$$

  2. Optimize the log-likelihood function by taking the derivatives with respect to the parameter of interest.

     Set to zero and solve for the parameter of interest.

     $$\ell'(\theta) = \frac{d}{d\theta}\ell(\theta) = 0 \qquad \rightarrow \qquad \hat{\theta} = \text{potential MLE}$$

  3. Verify that the global maximum of the log-likelihood function occurs at $\theta = \hat{\theta}$.

     Find the second derivative of the log-likelihood function, then plug in $\hat{\theta}$ and see if less than zero.

     $$\ell''(\theta) = \frac{d^2}{d\theta^2}\,\ell(\theta) \qquad \rightarrow \qquad \ell''(\hat{\theta}) \overset{?}{<} 0$$

     If so, then we have $\hat{\theta}_{MLE}$.

- Finding MLEs for functions of parameters

  Invariance property of MLEs: If $\hat{\theta}$ is the MLE of θ, then for any function $\tau(\theta)$, the MLE of $\tau(\theta)$ is $\tau(\hat{\theta})$

Distributions

Discrete Distributions	
Discrete uniform (N_0, N_1)	
Pmf	$P(X = x \mid N_0, N_1) = \frac{1}{N_1 - N_0 + 1}; \qquad x = N_0, \ldots, N_1; \qquad N_0 \leq N_1$
Mean and Variance	$E(X) = \frac{N_0 + N_1}{2}, \qquad V(X) = \frac{(N_1 - N_0 + 1)^2 - 1}{12}$
Mgf	$M_X(t) = \frac{1}{N_1 - N_0 + 1} \sum_{x=N_0}^{N_1} \mathrm{e}^{tx}$
Notes	
Bernoulli (p)	
Pmf	$P(X = x \mid p) = p^x (1-p)^{1-x}; \qquad x = 0, 1; \qquad 0 < \mathrm{p} < 1$
Mean and Variance	$E(X) = p, \qquad V(X) = p(1-p) = pq$
Mgf	$M_X(t) = (1-p) + p\mathrm{e}^t = q + p\mathrm{e}^t$
Notes	Special case of binomial with $n = 1$.
Binomial (n, p)	
Pmf	$P(X = x \mid n, p) = \binom{n}{x} p^x (1-p)^{n-x}; \qquad x = 0, 1, \ldots, n; \qquad 0 < p < 1$
Mean and Variance	$E(X) = np, \qquad V(X) = np(1-p) = npq$
Mgf	$M_X(t) = (q + p\mathrm{e}^t)^n$
Notes	Sum of *iid* bernoulli RVs.
Geometric (p)	
Pmf	$P(X = x \mid p) = q^{x-1} p; \qquad x = 1, 2, \ldots; \qquad 0 < p < 1$
Cdf	$F_X(x \mid p) = 1 - q^x$
Mean and Variance	$E(X) = \frac{1}{p}, \qquad V(X) = \frac{1-p}{p^2} = \frac{q}{p^2}$
Mgf	$M_X(t) = \frac{p\mathrm{e}^t}{1 - q\mathrm{e}^t}; \qquad t < -\ln(q)$
Notes	Special case of negative binomial with $r = 1$. * See other geometric probabilities. Alternate form $Y = X - 1$. This distribution is *memoryless*: $P(X > s \mid X > t) = P(X > s - t); \qquad s > t$.

Negative binomial (r,p)

Pmf	$P(X=x \mid r,p) = P(X=x \mid r,p) = \binom{x-1}{r-1} p^r q^{x-r}; \qquad x = r, r+1, \ldots; \qquad 0<p<1$
Mean and Variance	$E(X) = \frac{r}{p}, \qquad V(X) = \frac{r(1-p)}{p^2} = \frac{rq}{p^2}$
Mgf	$M_X(t) = \left[\frac{pe^t}{1-qe^t}\right]^r; \qquad t < -\ln(q)$
Notes	Sum of *iid* geometric RVs.

Hypergeometric (N, M, K)

Pmf	$P(X=x \mid r,p) = P(X=x \mid N,M,K) = \frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}}; \qquad x = 0,1,\ldots,\min(M,K)$
Mean and Variance	$E(X) = K\left(\frac{M}{N}\right), \qquad V(X) = K\left(\frac{M}{N}\right)\left(\frac{N-M}{N}\right)\left(\frac{N-K}{N-1}\right)$
Mgf	
Notes	If do not require $M \geq K$, $\mathcal{X} = \{\max(0, K+M-N), \ldots, \min(M,K)\}$, mean and variance converge to that of binomial $(n=K, p=M/K)$ when $N \to \infty$.

Poisson (λ)

Pmf	$P(X=x \mid \lambda) = \frac{e^{-\lambda}\lambda^x}{x!}, \qquad x = 0,1,2,\ldots; \qquad \lambda > 0$
Mean and Variance	$E(X) = \lambda, \qquad V(X) = \lambda$
Mgf	$M_X(t) = e^{\lambda(e^t-1)}$
Notes	If $X_i \overset{\perp\!\!\!\perp}{\sim}$ Poisson (λ_i), then $\sum X_i \sim \text{Poisson}(\lambda = \sum \lambda_i)$.

Other geometric probabilities

- Let $X \sim \text{Geometric}(p)$.

$$\begin{aligned}
P(X < \infty) &= 1 \\
P(X > x) &= q^x \\
P(X \geq x) &= q^{x-1} \\
P(a < X \leq b) &= q^a - q^b \\
P(a \leq X \leq b) &= q^{a-1} - q^b
\end{aligned}$$

Continuous Distributions

Continuous uniform (a, b)

Pdf	$f(x \mid a, b) = \frac{1}{b-a}, \qquad a \le x \le b; \qquad a, b \in \mathbb{R}, \qquad a \le b$
Cdf	$F(x) = \frac{x-a}{b-a} \qquad a \le x \le b$
Survival	$S(t) = \frac{b-t}{b-a} \qquad a \le t \le b$ if $T \sim \text{Uniform}\,(a, b)$
Mean and Variance	$E(X) = \frac{a+b}{2}; \qquad V(X) = \frac{(b-a)^2}{12}$
Mgf	$M_X(t) = \frac{e^{tb} - e^{ta}}{t(b-a)} \qquad t \neq 0$
Notes	

Exponential (λ)

Pdf	$f(t \mid \lambda) = \lambda e^{-\lambda t}, \qquad t \ge 0; \qquad \lambda > 0$
Cdf	$F(t) = 1 - e^{-\lambda t} \qquad t \ge 0$
Survival	$S(t) = e^{-\lambda t} \qquad t \ge 0$
Mean and Variance	$E(X) = \frac{1}{\lambda}; \qquad V(X) = \frac{1}{\lambda^2}$
Mgf	$M_X(t) = \frac{\beta}{\beta - t} \qquad t < \beta;$ if $T \sim \text{Exp}\,(\beta)$
Notes	Special case of gamma with $\alpha = 1, \beta$. This distribution is *memoryless*: $P(T > a + b \mid T > a) = P(T > b); \qquad a, b > 0$. Alternate parameterization is with scale $\theta = 1/\lambda$.

Gamma (α, β)

Pdf	$f(x \mid \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\, x^{\alpha-1}\, e^{-\beta x}, \qquad x \ge 0; \qquad \alpha, \beta > 0$
Cdf	N/A
Mean and Variance	$E(X) = \frac{\alpha}{\beta} \qquad V(X) = \frac{\alpha}{\beta^2}$
Mgf	$M_X(t) = \left(\frac{\beta}{\beta - t}\right)^\alpha \qquad t < \beta$
Notes	Sum of *iid* exponential RVs. A special case is exponential $(\alpha = 1, \beta)$. Alternate parameterization is with scale $\theta = 1/\beta$.

Normal (μ, σ^2)

Pdf	$f(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \qquad -\infty < x < \infty; \qquad -\infty < \mu < \infty, \quad \sigma > 0$
Cdf	N/A
Mean and Variance	$E(X) = \mu, \qquad V(X) = \sigma^2$
Mgf	$M_X(t) = \exp\left[\mu t + \frac{\sigma^2 t^2}{2}\right]$
Notes	Special case: Standard normal $Z \sim \text{Normal}\,(\mu = 0, \sigma^2 = 1)$.

Lognormal (μ, σ^2)

Pdf	$f(y \mid \mu, \sigma^2) = \frac{1}{y\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(\ln(y)-\mu)^2}{2\sigma^2}\right]; \qquad y \geq 0; \qquad -\infty < \mu < \infty; \qquad \sigma > 0$
Mean and Variance	$E(Y) = \mathrm{e}^{\mu + \frac{\sigma^2}{2}}, \qquad V(Y) = \mathrm{e}^{2\mu+\sigma^2}(\mathrm{e}^{\sigma^2} - 1)$
Mgf	
Notes	If $Y \sim$ Lognormal $\Longrightarrow \ln(Y) \sim$ Normal (μ, σ^2); equivalently, if $X \sim$ Normal (μ, σ^2) and $Y = \mathrm{e}^X \Longrightarrow Y \sim$ Lognormal. μ and σ^2 represent the mean and variance of the normal random variable X which appears in the exponent.

Beta (α, β)

Pdf	$f(x \mid \alpha, \beta) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}; \qquad 0 \leq x \leq 1; \qquad \alpha, \beta > 0$
Mean and Variance	$E(X) = \frac{\alpha}{\alpha+\beta}, \qquad V(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$
Mgf	
Notes	$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\,\mathrm{d}x = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$

Chi-square, χ^2 (r)

Pdf	$f(x \mid r) = \frac{1}{\Gamma(\frac{r}{2})2^{r/2}} x^{\frac{r}{2}-1} \mathrm{e}^{-\frac{x}{2}}, \qquad x \geq 0; \qquad r = 0.5, 1, 1.5, 2, \ldots$
Cdf	N/A
Mean and Variance	$E(X) = r, \qquad V(X) = 2r$
Mgf	$M_X(t) = \left(\frac{\theta}{\theta - 2t}\right)^{r/2} \qquad t < 1/2$
Notes	Special case of (scale) gamma with $\alpha = r/2, \theta = 2$.

t (r)

Pdf	$f(t \mid r) = f_T(t) = \frac{\Gamma(\frac{r+1}{2})}{\frac{1}{\sqrt{r\pi}}\Gamma(\frac{r}{2})}\left(\frac{1}{(1+t^2/r)^{(r+1)/2}}\right), \qquad -\infty < t < \infty$
Cdf	N/A
Mean and Variance	$E(T) = 0 \quad$ if $r > 1, \qquad V(X) = \frac{r}{r-2} \quad$ if $r > 2$
Mgf	N/A
Notes	See derivation notes above.

F (r_1, r_2)

Notes	See derivation notes above.