MATH 324 Statistics - Midterm Crib Sheet

Expectation: $\mathbb{E}[g(Y)] = \sum_y g(y)p(y)$ discrete, $\int_{-\infty}^{\infty} g(y)f(y)\mathrm{d}y$ continuous; $\mathbb{E}[\overline{Y}^2] = \mu^2 + \frac{\sigma^2}{n}$
Variance:
$\mathbb{V}[Y] = \mathbb{E}[(Y-\mathbb{E}[Y])^2] = \mathbb{E}[Y^2] - \mathbb{E}[Y]^2$, $\mathbb{V}[c] = 0$, $\mathbb{V}[cY] = c^2\mathbb{V}[Y]$, $\mathbb{V}[X+Y] = \mathbb{V}[X] + 2\mathrm{Cov}[X,Y] + \mathbb{V}[Y] = \mathbb{V}[X] + \mathbb{V}[Y]$ if X and Y indep, $\mathbb{V}[aX+bY+c] = a^2\mathbb{V}[X] + 2ab\mathrm{Cov}[X,Y] + b^2\mathbb{V}[Y]$
$\mathbb{V}[\overline{Y}] = 1/n^2\mathbb{V}[\sum Y_i] = 1/n^2\left(\sum\mathbb{V}[Y_i] + 2\sum\sum_{1\le i\le j\le n}\mathrm{Cov}(Y_i,Y_j)\right) = 1/n^2\sum\mathbb{V}[Y_i] = n\sigma^2/n^2 = \sigma^2/n$
Covariance and correlation:
$\mathrm{Cov}[Y_1,Y_2] = \mathbb{E}[Y_1Y_2] - \mathbb{E}[Y_1]\mathbb{E}[Y_2]$, $\quad$ $\mathrm{Corr}[Y_1,Y_2] = \frac{\mathrm{Cov}[Y_1,Y_2]}{\sqrt{\mathbb{V}[Y_1]\mathbb{V}[Y_2]}}$
Joint pdf symmetric on Y_1 and $Y_2 \Rightarrow$ same marginal distribs & $\mathbb{E} \Rightarrow$ no correlation.
Standard deviation of $Y = \sqrt{\mathbb{V}[Y]}$
Sample std deviation: $s = \sqrt{s^2} = \sqrt{\sum(Y-\overline{Y})^2/(n-1)}$; sample variance s^2 unbiased estimator of σ^2
MSE($\hat{\theta}$): $\mathbb{E}[(\hat{\theta}-\theta)^2] = \mathbb{E}[\hat{\theta}^2] - 2\theta\mathbb{E}[\hat{\theta}] + \theta^2$; $\mathbb{V}[\hat{\theta}] = \mathrm{MSE}[\hat{\theta}]$ if $\hat{\theta}$ unbiased estimator of θ
Expected vals and std errors of some common point estimators

target param θ	sample size(s)	point estimator $\hat{\theta}$	$\mathbb{E}[\hat{\theta}]$	std error $\sigma_{\hat{\theta}}$
μ	n	$\overline{Y}$	μ	$\frac{\sigma}{\sqrt{n}}$
p	n	$\hat{p} = Y/n$	p	$\sqrt{\frac{pq}{n}}$
$\mu_1 - \mu_2$	n_1 and n_2	$\overline{Y}_1 - \overline{Y}_2$	$\mu_1 - \mu_2$	$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$
$p_1 - p_2$	n_1 and n_2	$\hat{p}_1 - \hat{p}_2$	$p_1 - p_2$	$\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}$

2-standard-error: $2\sigma_{\hat{\theta}}$
$100(1-\alpha)\%$ **confidence interval:**
$P(a \le U \le b) = 1-\alpha$, $P(U \le a) = \int_0^a f_U(u)\mathrm{d}u = P(U \ge b) = \int_b^{\infty} f_U(u)\mathrm{d}u = \alpha/2$
In **large samples**, estimators have **normal sampling distributions:** $\mathrm{CI} = \hat{\theta} \pm z_{\alpha/2}\sigma_{\hat{\theta}}$, $\quad z_{\alpha/2}$ critical value
Small-sample CIs:
for μ: $\mathrm{CI} = \overline{Y} \pm t_{\alpha/2}(S/\sqrt{n})$, $\nu = df = n-1$
for $\mu_1-\mu_2$: $\mathrm{CI} = \overline{Y}_1 - \overline{Y}_2 \pm t_{\alpha/2}S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$, $\nu = n_1+n_2-2$ and **pooled sample estimator/variance** $S_p^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$;
S_p^2 **unbiased and consistent** est. of σ^2, $\mathbb{E}[S_p^2] = \sigma^2$, $\mathbb{V}[S_p^2] = \frac{2\sigma^4}{n_1+n_2-2} = \frac{\sigma^4}{n-1}$
CI for σ^2: $\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2}}, \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}}\right)$, $\nu = n$
Choosing sample size: Solve $z_{\alpha/2}\sigma_{\hat{\theta}} = B$ for n, with B = desired bound
Show U pivotal qty: Show $F_U(u)$ indep of θ, e.g. $U = Y/\theta$, $F_U(u) = P(U \le u) = P(Y \le \theta u) = 2u - u^2 \quad 0 < u < 1$, indep of θ
Cdf to pdf: $F_Y(y) = \int_0^y f_Y(t)\mathrm{d}t$
Order statistics: $f_{(n)}(y) = nF(y)^{n-1}f(y)$
Efficiency: $\mathrm{eff}(\hat{\theta}_1,\hat{\theta}_2) = \mathbb{V}[\hat{\theta}_2]/\mathbb{V}[\hat{\theta}_1]$
Consistency:
$\hat{\theta}$ unbiased and $\lim_{n\to\infty}\mathbb{V}[\hat{\theta}] = 0$
$\hat{\theta} \to \theta$ and $\hat{\theta}' \to \theta' \Rightarrow \hat{\theta} \pm \hat{\theta}' \to \theta \pm \theta', \hat{\theta} \times \hat{\theta}' \to \theta \times \theta', \hat{\theta}/\hat{\theta}' \to \theta/\theta'$ if $\theta' \neq 0$, $f(\hat{\theta})(\theta)$ if f real-valued fn continuous at θ
Central Limit Theorem: $U_n = \frac{\sum_i^n Y - n\mathbb{E}[Y]}{\sqrt{\mathbb{V}[Y]n}} = \frac{\sum_i^n Y - n\mu}{\sigma\sqrt{n}} = \frac{\overline{Y}-\mathbb{E}[Y]}{\sqrt{\mathbb{V}[Y]/n}} \to$ std Normal distr.. $W_n \to 1 \Rightarrow U_n/W_n \to$ std Normal distr.
Sufficiency: U sufficient for θ if
$P(Y_1 = y_1, ..., Y_n = y_n | U = u) = P(Y_1 = y_1, ..., Y_n = y_n, U = u)/P(U = u)$ indep of θ
Factorization criterion proving sufficiency:
$L(\theta) = f$ or $p(y_1,...,y_n|\theta) = \prod_i^n f(y_i|\theta)$ if iid $= g(u,\theta) \times h(y_1,...,y_n)$, g fn of only u and θ, h not fn of θ
Rao-Blackwell Theorem: $\hat{\theta}$ unbiased est. of θ and $\mathbb{V}[\hat{\theta}] < \infty$ and U sufficient stat. for $\theta \Rightarrow \hat{\theta}^* = \mathbb{E}[\hat{\theta}|U]$, $\mathbb{E}[\hat{\theta}^*] = \theta, \mathbb{V}[\hat{\theta}^*] < \mathbb{V}[\hat{\theta}]$
Typically $\hat{\theta}^*$ is **MVUE** of θ
Minimum variance unbiased estimation (MVUE): Some fn of sufficient U, $h(U)$, $\mathbb{E}[h(U)] = \theta \Rightarrow h(U)$ MVUE of θ
MLE: Solve $\frac{\partial L(\theta)}{\partial\theta} = 0$ for θ
Law of Large Numbers: Sample vals converges to theoretical vals e.g. $\overline{Y} \to \mathbb{E}[Y]$ **Bernoulli(p):** $p(y) = p^y(1-p)^{1-y}$, $\mathbb{E}[Y] = p$, $\mathbb{V}[Y] = p(1-p) \qquad y = 0, 1$
Binomial(n,p): $(a+b)^n = \sum_{j=0}^n \binom{n}{j}a^jb^{n-j}$, $p(y) = \binom{n}{y}p^y(1-p)^{n-y}$, $\mathbb{E}[Y] = np$, $\mathbb{V}[Y] = np(1-p) \qquad y = 0:n$
Poisson(λ): $p(y) = \frac{\lambda^y}{y!}e^{-\lambda}$, $\mathbb{E}[Y] = \mathbb{V}[Y] = \lambda \qquad y = 0, 1, ...; \quad m(t) = \exp\{\lambda(e^t-1)\}$
$\sum_i^n Y_i \sim \mathrm{Poisson}(n\lambda)$
Power family(α,θ): $f(y) = \alpha y^{\alpha-1}/\theta^\alpha \quad 0 \le y \le \theta, \quad$ 0 otherwise, $F(y) = \frac{y^\alpha}{\theta^\alpha}$, $\mathbb{E}[Y] = \alpha\theta/(\alpha+1)$
Uniform(θ_1,θ_2): $f(y) = \frac{1}{\theta_2-\theta_1} \quad y \in (\theta_1,\theta_2), \quad$ 0 otherwise, $F(y) = \frac{y-\theta_1}{\theta_2-\theta_1}$, $\mathbb{E}[Y] = \frac{\theta_1+\theta_2}{2}$, $\mathbb{V}[Y] = \frac{(\theta_2+\theta_1)^2}{12}$
Gamma(α,β): $f(y) = \frac{1}{\beta^\alpha\Gamma(\alpha)}y^{\alpha-1}e^{-y/\beta} \qquad y \ge 0, \quad$ 0 otherwise
$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1}e^{-t}\mathrm{d}t, \qquad \Gamma(n) = (n-1)\Gamma(n-1) = (n-1)!, \qquad \Gamma(1/2) = \sqrt{\pi}$
$\mathbb{E}[Y] = \alpha\beta, \qquad \mathbb{V}[Y] = \alpha\beta^2, \qquad \mathbb{E}[Y^2] = \alpha(\alpha+1)\beta^2$, $\mathbb{E}[Y^3] = \alpha(\alpha+1)(\alpha+2)\beta^2$, ... $\qquad m(t) = \left(\frac{1}{1-\beta t}\right)^\alpha$

Chi-squared(ν): $\alpha = \nu/2$, $\beta = 2$, $\nu = 1, 2, ...$

Exponential(β): $\alpha = 1$, standard $\beta = 1$, $f(y) = \frac{1}{\beta}e^{y/\beta}$, $F(y) = 1 - e^{-y/\beta}$, $\mathbb{E}[Y] = \beta$, $\mathbb{V}[Y] = \beta^2$

Normal(μ, σ^2): $f(y) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{1}{2\sigma^2}(y-\mu)^2\right\}$, $F(y) = \int_{-\infty}^{y} f(t)\mathrm{d}t$; $\mathbb{E}[Y] = \mu$, $\mathbb{V}[Y] = \sigma^2$

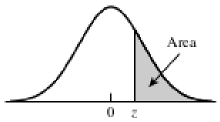
Standard Normal(μ, σ^2): $\mu = 0$, $\sigma = 1$; $f(y) = \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{y^2}{2}\right\}$

Beta: $f(y) = \frac{1}{B(\alpha,\beta)}y^{\alpha-1}(1-y)^{\beta-1}$, $y \in [0,1]$, 0 otherwise; $B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} = \int_0^1 y^{\alpha-1}(1-y)^{\beta-1}\mathrm{d}y$

$\alpha = \beta = 1 \Rightarrow Y \sim$ Uniform(0, 1)

Weibull(α, β): $f(y) = \alpha\beta y^{\alpha-1}e^{-\beta y^{\alpha}}$ $y \geq 0$, 0 otherwise, $F(y) = 1 - e^{-\beta y^{\alpha}}$, $\mathbb{E}[Y] = \frac{\Gamma(1+1/\alpha)}{\beta^{1/\alpha}}$, $\mathbb{V}[Y] = \frac{\Gamma(1+2/\alpha)-\Gamma(1+1/\alpha)^2}{\beta^{2/\alpha}}$

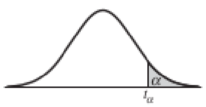
Table 4 Normal Curve Areas
Standard normal probability in right-hand tail
(for negative values of z, areas are found by symmetry)



	Second decimal place of z									
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0722	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0352	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0017	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.00135									
3.5	.000 233									
4.0	.000 031 7									
4.5	.000 003 40									
5.0	.000 000 287									

From R. E. Walpole, *Introduction to Statistics* (New York: Macmillan, 1968).

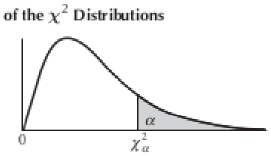
Table 5 Percentage Points of the t Distributions



$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$	df
3.078	6.314	12.706	31.821	63.657	1
1.886	2.920	4.303	6.965	9.925	2
1.638	2.353	3.182	4.541	5.841	3
1.533	2.132	2.776	3.747	4.604	4
1.476	2.015	2.571	3.365	4.032	5
1.440	1.943	2.447	3.143	3.707	6
1.415	1.895	2.365	2.998	3.499	7
1.397	1.860	2.306	2.896	3.355	8
1.383	1.833	2.262	2.821	3.250	9
1.372	1.812	2.228	2.764	3.169	10
1.363	1.796	2.201	2.718	3.106	11
1.356	1.782	2.179	2.681	3.055	12
1.350	1.771	2.160	2.650	3.012	13
1.345	1.761	2.145	2.624	2.977	14
1.341	1.753	2.131	2.602	2.947	15
1.337	1.746	2.120	2.583	2.921	16
1.333	1.740	2.110	2.567	2.898	17
1.330	1.734	2.101	2.552	2.878	18
1.328	1.729	2.093	2.539	2.861	19
1.325	1.725	2.086	2.528	2.845	20
1.323	1.721	2.080	2.518	2.831	21
1.321	1.717	2.074	2.508	2.819	22
1.319	1.714	2.069	2.500	2.807	23
1.318	1.711	2.064	2.492	2.797	24
1.316	1.708	2.060	2.485	2.787	25
1.315	1.706	2.056	2.479	2.779	26
1.314	1.703	2.052	2.473	2.771	27
1.313	1.701	2.048	2.467	2.763	28
1.311	1.699	2.045	2.462	2.756	29
1.282	1.645	1.960	2.326	2.576	inf.

From "Table of Percentage Points of the t-Distribution." Computed by Maxine Merrington, *Biometrika*, Vol. 32 (1941), p. 300.

Table 6 Percentage Points of the χ^2 Distributions



df	$\chi^2_{0.995}$	$\chi^2_{0.990}$	$\chi^2_{0.975}$	$\chi^2_{0.950}$	$\chi^2_{0.900}$
1	0.0000393	0.0001571	0.0009821	0.0039321	0.0157908
2	0.0100251	0.0201007	0.0506356	0.102587	0.210720
3	0.0717212	0.114832	0.215795	0.351846	0.584375
4	0.206990	0.297110	0.484419	0.710721	1.063623
5	0.411740	0.554300	0.831211	1.145476	1.61031
6	0.675727	0.872085	1.237347	1.63539	2.20413
7	0.989265	1.239043	1.68987	2.16735	2.83311
8	1.344419	1.646482	2.17973	2.73264	3.48954
9	1.734926	2.087912	2.70039	3.32511	4.16816
10	2.15585	2.55821	3.24697	3.94030	4.86518

Table 6 (*Continued*)

$\chi^2_{0.100}$	$\chi^2_{0.050}$	$\chi^2_{0.025}$	$\chi^2_{0.010}$	$\chi^2_{0.005}$	df
2.70554	3.84146	5.02389	6.63490	7.87944	1
4.60517	5.99147	7.37776	9.21034	10.5966	2
6.25139	7.81473	9.34840	11.3449	12.8381	3
7.77944	9.48773	11.1433	13.2767	14.8602	4
9.23635	11.0705	12.8325	15.0863	16.7496	5
10.6446	12.5916	14.4494	16.8119	18.5476	6
12.0170	14.0671	16.0128	18.4753	20.2777	7
13.3616	15.5073	17.5346	20.0902	21.9550	8
14.6837	16.9190	19.0228	21.6660	23.5893	9
15.9871	18.3070	20.4831	23.2093	25.1882	10