STA9715 - Test 3 - Formula Sheet

Advanced Inequalities

- Chernoff - Gaussian: if $X \sim \mathcal{N}(\mu, \sigma^2)$, then $\mathbb{P}(|X-\mu| > t) \leq 2e^{-t^2/2\sigma^2}$ and $\mathbb{P}(X > \mu + t) \leq e^{-t^2/2\sigma^2}$
- Chernoff - Bounded: if X takes values in the range $[a, b]$ with mean μ, then $\mathbb{P}(|X-\mu| > t) \leq 2e^{-2t^2/(b-a)^2}$ and $\mathbb{P}(X > \mu + t) \leq e^{-2t^2/(b-a)^2}$. For means, $\mathbb{P}(|\overline{X}_n - \mu| > t) \leq 2e^{-2nt^2/(b-a)^2}$ and $\mathbb{P}(\overline{X}_n \geq \mu + t) \leq e^{-2nt^2/(b-a)^2}$

Moment Generating Functions

- Moment Generating Function: $\mathbb{M}_X(t) = \mathbb{E}[e^{tX}]$
- MGF to Moments: $\mathbb{E}[X^k] = \mathbb{M}_X^{(k)}(0)$
- MGF of linear transforms: $\mathbb{M}_{aX+b} = \mathbb{E}[e^{t(aX+b)}] = e^{tb}\mathbb{M}_X(ta)$
- MGF of sums of independent RVs: $\mathbb{M}_{X+Y}(t) = \mathbb{M}_X(t)\mathbb{M}_Y(t)$
- If $\mathbb{M}_X(t) = \mathbb{M}_Y(t)$, then X and Y have the same distribution.

Limit Theory

- Convergence in Probability: $X_n \xrightarrow{P} X_*$ means $\mathbb{P}(|X_n - X_*| > \epsilon) \xrightarrow{n\to\infty} 0$
- Convergence in Distribution: $X_n \xrightarrow{d} X_*$ means $\mathbb{E}[f(X_n)] \xrightarrow{n\to\infty} \mathbb{E}[f(X_*)]$ for 'reasonable' functions $f(\cdot)$.
- Law of Large Numbers: if $X_1, X_2, \dots$ are IID random variables with mean μ and finite variance, $\overline{X}_n \xrightarrow{P} \mu$ for $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$.
- Central Limit Theorem: if $X_1, X_2, \dots$ are IID random variables with mean μ and finite variance σ^2, then $\sqrt{n}\left(\frac{\overline{X}_n - \mu}{\sigma}\right) \xrightarrow{d} \mathcal{N}(0, 1)$. More usefully: $\overline{X}_n \xrightarrow{d} \mathcal{N}(\mu, \sigma^2/n)$
- Delta Method: if $X_n \xrightarrow{d} \mathcal{N}(\mu, \sigma^2)$, then $g(X_n) \xrightarrow{d} \mathcal{N}(g(\mu), \sigma^2 g'(\mu)^2)$ for any differentiable $g(\cdot)$
- Glivenko-Cantelli Theorem ('Fundamental Theorem of Statistics'): the sample CDF $\hat{F}_n(\cdot)$ converges to the true CDF $F_X(\cdot)$ at all points where $F_X(\cdot)$ is continuous
- Massart's Inequality: $\mathbb{P}(\max_x |\hat{F}_n(x) - F_X(x)| > \epsilon) \leq 2e^{-2n\epsilon^2}$ for any distribution and any $\epsilon > 0$

Key Statistical Distributions

- Gaussian / Normal - See above. Standardizing, CLT, Delta Method.
- χ^2_k sum of squares of k IID Standard Normals. Arises from 'goodness of fit' type statistics (*e.g.*, SSE in OLS)
- t_k - (Student's) t distribution with k degrees of freedom. $t_k \stackrel{d}{=} Z/\sqrt{\chi^2_k/k}$ where $Z \perp\!\!\!\perp \chi^2_k$. t_1 is a Cauchy; t_∞ is a standard normal. Can replace Z with other normal distributions. Arises in testing with unknown variance.
- χ^2_2 is an Expo(1/2) distribution with mean 2
- Gamma distribution: sum of n exponential distributions with mean $1/\theta$ is $\Gamma(n, \theta)$ distributed. $\Gamma(n/2, 2) \stackrel{d}{\to} \chi^2_n$
- Beta distribution = Gamma ratio. $X \sim \Gamma(\alpha, \theta), Y \sim \Gamma(\beta, \theta) \implies X/(X+Y) \sim B(\alpha, \beta)$. Support on $[0, 1]$
- F distribution: $\frac{\chi^2_{k_1}/k_1}{\chi^2_{k_2}/k_2}$

Name	Parameters	Density	Mean	Variance	MGF
Standard Normal	None	$\phi(z) = e^{-z^2/2}/\sqrt{2\pi}$	0	1	$e^{t^2/2}$
Normal	Mean μ, StdDev σ	$e^{-(x-\mu)^2/2\sigma^2}/\sqrt{2\pi\sigma^2}$	μ	σ^2	$e^{\mu t + \sigma^2 t^2/2}$
χ^2	Degrees of freedom k	$\propto x^{k/2-1}e^{-x/2}$	k	$2k$	$(1-2t)^{-k/2}$
Standard Student's t	Degrees of freedom k	$\propto (1 + x^2/k)^{-(k+1)/2}$	0	$k/(k-2)$	NA
Gamma	Shape k, Scale θ	$\propto x^{k-1}e^{-x/\theta}$	$k\theta$	$k\theta^2$	$(1-\theta t)^{-k}$
Beta	Shapes α, β	$\propto x^{\alpha-1}(1-x)^{\beta-1}$	$\alpha/(\alpha+\beta)$	$(\alpha\beta)(\alpha+\beta)^{-2}(\alpha+\beta+1)^{-1}$	Hard
F	Deg. Freedom k_1, k_2	Hard	$k_2/(k_2-2)$	Hard	NA

Sampling (Probability Integral Transform): If X has CDF F_X, $F_X^{-1}(U) \stackrel{d}{=} X$ for $U \sim \mathcal{U}([0, 1])$

Sampling (Box-Mueller): Let $R^2 \sim \chi^2_2$ and $\Theta \sim \mathcal{U}([0, 2\pi])$; then $X = R\cos\Theta, Y = R\sin\Theta$ are independent Z_1, Z_2