Mathematics 333 : Probability and Statistics
Formula Sheet

1. Sample variance $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$

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

3. (*Addition Rule for Probabilities.*) For events A and B,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

4. For positive integers n and k, with $k \leq n$,

i) Permutations: $P_k^n = n(n-1)\cdots(n-k+1) = \frac{n!}{(n-k)!}$

ii) Combinations: $\binom{n}{k} = \frac{n!}{k!(n-k)!}$

5. Conditional probability formula: $P(A \mid B) = \frac{P(A \cap B)}{P(B)}$

6. (*Independence*) Events A and B are independent if and only if:

$$P(A \cap B) = P(A)P(B)$$

7. (*Bayes' formula*) If $\{A_1, A_2, \cdots, A_n\}$ are mutually exclusive and exhaustive events, and B is any event, then

$$P(A_k \mid B) = \frac{P(A_k)P(B \mid A_k)}{\sum_{i=1}^{n} P(A_i)P(B \mid A_i)}, \qquad k = 1, 2, \cdots, n$$

8. The c.d.f. $F(x)$ of a random variable X :

$$F(x) = P(X \leq x) = \begin{cases} \sum_{y:y \leq x} f(y), & \text{for } X \text{ discrete} \\ \int_{-\infty}^{x} f(y)\,dy, & \text{for } X \text{ continuous} \end{cases}$$

9. For any real valued function $h(x)$,

$$E\{h(X)\} = \begin{cases} \sum_x h(x)f(x), & \text{for } X \text{ discrete} \\ \int_{-\infty}^{\infty} h(x)f(x)\,dx, & \text{for } X \text{ continuous} \end{cases}$$

10. binomial$(x; n, p) = \binom{n}{x} p^x (1-p)^{n-x}$, $x = 0, 1, 2, \cdots, n$;
$E(X) = np$, $V(X) = np(1-p)$.

11. Poisson$(x;\lambda) = e^{-\lambda}\left(\frac{\lambda^x}{x!}\right), x = 0,1,2,\cdots;\quad E(X) = V(X) = \lambda.$
12. Exponential density: $f(x) = \lambda e^{-\lambda x}$, $x > 0; EX = \lambda^{-1}$, $V(X) = \lambda^{-2}$.
13. Z-score of the sample mean $\overline{X}$ for a sample of size n, is $Z = \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$.

14. (i) A *large sample* $100(1-\alpha)\%$ confidence interval (CI) for the population mean μ, is:

$$\overline{X} \pm z_{\frac{\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right)$$

(ii) A *small sample* $100(1-\alpha)\%$ confidence interval (CI) for a *normal* population mean μ, is:

$$\overline{X} \pm t_{\frac{\alpha}{2},n-1}\left(\frac{s}{\sqrt{n}}\right)$$

(iii) A *large sample* $100(1-\alpha)\%$ confidence interval (CI) for the population proportion p is:

$$\hat{p} \pm z_{\frac{\alpha}{2}}\left(\sqrt{\frac{\hat{p}\hat{q}}{n}}\right)$$

15. A $100(1-\alpha)\%$ CI for the population variance σ^2 in a Normal distribution is :

$$\left(\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2},n-1}}\ ,\ \frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2},n-1}}\right)$$

16. An *exact* small sample $100(1-\alpha)\%$ CI for the difference $(\mu_1-\mu_2)$ of two normal means, with equal variances (unknown) is:

$$\overline{X} - \overline{Y} \pm t_{\frac{\alpha}{2},n+m-2}\left(s_p\sqrt{\frac{1}{m}+\frac{1}{n}}\right),$$

where $s_p^2 = \frac{(m-1)s_1^2+(n-1)s_2^2}{m+n-2}$, and $(\overline{X}, s_1^2)$, $(\overline{Y}, s_2^2)$ are the sample means and variances based on m and n random sample observations drawn from independent Normal populations with a common unknown variance.
17. A small sample *approximate* $100(1-\alpha)\%$ CI for the difference $(\mu_1-\mu_2)$ of two Normal means and possibly unequal variances, is:

$$(\bar{X}-\bar{Y}) \pm t_{\frac{\alpha}{2},\nu}\sqrt{\frac{s_1^2}{m}+\frac{s_2^2}{n}}$$

where, $\nu = \frac{\left(\{s_1^2/m\}+\{s_2^2/n\}\right)^2}{\frac{(s_1^2/m)^2}{m-1}+\frac{(s_2^2/n)^2}{n-1}}$ is the estimated degrees of freedom rounded down to the nearest integer.